NOTE

This Microfiche meets the quality requirements of the American National Standard for Microfiche of Documents (ANSI PH5.9-1975). However, because of imperfections in the document from which it was made, it contains images that are illegible or difficult to read. The source document for this microfiche is the best copy that could be obtained prior to filming.

Congressional Information Service, Inc.
Washington, D.C.
CIS

Congress of the United States
Office of Technology Assessment

ALTERNATIVES TO ANIMAL USE IN RESEARCH, TESTING, AND EDUCATION

Contractor Documents

Volume III: Testing and Economics

February 1986

"Animal Testing for Safety and Effectiveness," Thomas D. Sabourin, Betsy D. Carlton, Robin T. Faulk and L. Barry Goss, Environmental and Health Sciences Division, Battelle Columbus Laboratories, Columbus, Ohio.

"Animal Testing and Alternatives: Review of Animal Testing in FDA and EPA; Interrelationship to Products Liability and Uses of Alternatives to Animal Testing," Bruce F. Mackler, Meyer, Faller and Weisman, P.C., Washington, D.C.

"Economic and Policy Considerations," Henry R. Hertzfeld and Thomas D. Myers, Washington, D.C.

These are contractor documents that were used in preparing OTA's final Assessment Report. OTA makes these contractor documents available for the use of readers desiring a more detailed or technical discussion of an issue than can normally be accommodated in our final Report. As OTA contractor documents, they have not been reviewed or approved by the Technology Assessment Board. The findings and conclusions expressed are those of the authors and do not necessarily reflect the views of OTA, the Advisory Panel or the Technology Assessment Board.

ANIMAL TESTING FOR SAFETY AND EFFECTIVENESS

FINAL REPORT
January 29, 1985

Prepared By

Battelle's Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

Written By

Thomas D. Sabourin
Betsy D. Carlton
Robin T. Faulk
L. Barry Goss

Prepared For

U.S. Congress
Office of Technology Assessment
Contract No. 433-71400/BCL No. G-0408

TABLE OF CONTENTS

Page

DOMESTIC REGULATORY REQUIREMENTS 1

DATA SUBMISSION REQUIREMENTS, METHODOLOGIES AND PROTOCOLS 1

Consumer Product Safety Commission 1

Consumer Product Safety Act/Federal Hazardous Substances Act 1
Poison Prevention Packaging Act 2
Lead-Based Paint Poison Prevention Act 3

Department of Agriculture 3

Egg Products Inspection Act 3
Poultry Products Inspection Act 4
Federal Meat Inspection Act 4

Department of Health and Human Services 4

FOOD AND DRUG ADMINISTRATION 4

Federal Food, Drug, and Cosmetic Act (FFDCA) 4
Fair Packaging & Labeling Act 12

NATIONAL TOXICOLOGY PROGRAM 12
NATIONAL INSTITUTE OF OCCUPATIONAL SAFETY AND HEALTH 13

Occupational Safety and Health Act (OSHA) 13
Federal Mine Safety and Health Act 14

Department of Housing and Urban Development 14

Lead-Based Paint Poison Prevention Act 14

Department of Transportation 15

MATERIALS TRANSPORTATION BUREAU 15

Hazardous Material Transportation Act 15
COAST GUARD 16

Port and Waterways Safety Act 16

FEDERAL RAILWAY ADMINISTRATION 16

Federal Railway Safety Act 16

TABLE OF CONTENTS
(Continued)

Page

Environmental Protection Agency.. 16

Clean Water Act.. 16
Clean Air Act.. 19
Resource Conservation and Recovery Act......................... 20
Safe Drinking Water Act.. 21
Toxic Substances Control Act................................... 21
Federal Insecticide Fungicide and Rodenticide Act.............. 25

SUMMARY.. 36

FOREIGN REGULATORY REQUIREMENTS.. 36

DATA SUBMISSION REQUIREMENTS... 36

European Economic Community (EEC)...................................... 36
Non-EEC Countries.. 41

Japan.. 41

METHODOLOGIES AND PROTOCOLS.. 41

World Health Organization.. 41
Organization of Economic Cooperation and Development................... 46

STATE-OF-THE-ART OF DESCRIPTIVE AND MECHANISTIC TOXICOLOGY.............. 49

Descriptive Toxicology... 49

Acute Toxicity... 50
Short-Term Toxicity (14-28 days)....................................... 51
Subchronic Toxicity.. 52
Chronic Toxicity... 53
Special Studies.. 55

Mechanistic Toxicology... 55
Current Needs and Future Trends.. 62

Needs.. 62
Trends... 63

ALTERNATIVES TO ANIMAL USE IN TOXICITY TESTING.......................... 65

ALTERNATIVE TECHNOLOGIES... 65

Mammalian Approaches... 65

TABLE OF CONTENTS
(Continued)

Page

Modified Whole Animal Approaches........................ 65
In Vitro Culture Techniques............................ 67
Non-Invasive Human Methods............................. 70

Nonmammalian Approaches................................ 74

Vertebrates.. 74
Invertebrates.. 76
Microbes... 78
Plants... 79
Computer Models.. 80

INCORPORATION OF ALTERNATIVE TECHNOLOGIES INTO EXISTING TESTING REGIMENS.. 82

Validated Alternative Test Systems..................... 86

Alternative Toxicity Test Systems Under Development.... 86

POTENTIAL, CONSTRAINTS, AND CONCLUSIONS REGARDING THE INCORPORATION OF ALTERNATIVE TECHNOLOGIES.................. 93

POTENTIAL FOR INCORPORATING ALTERNATIVE TECHNOLOGIES.... 93

Advantages of Alternative Test Systems................. 93
Limitations of Alternative Tests Systems............... 94
Feasibility of Applying Alternative Test Systems....... 95

THE INTERFACE OF CHANGING SCIENCE AND REGULATORY POLICY. 96

Solutions.. 96

Modifying Precision.................................... 96
Redefining Methods..................................... 97
Redefining Test Criteria............................... 97
Reducing Developmental Projects........................ 98
Reducing Routine Testing Programs...................... 98
Replacing Existing Methodologies....................... 98

Changes Required to Accommodate........................

Alternative Technologies............................... 99

REFERENCES... 100

DOMESTIC REGULATORY REQUIREMENTS

Numerous legislative acts require biological testing for assessment of human health and safety. The specific animal testing requirements vary for each act and also vary within the individual acts depending upon factors such as intended use of the substance under consideration, projected production volume, potential for and extent of human exposure, and other factors. Most acts require mammalian testing for human health effects assessment, while some also require testing with nonmammalian models.

A summary of several of the major acts and their testing requirements under the jurisdiction of various federal agencies follows.

DATA SUBMISSION REQUIREMENTS, METHODOLOGIES AND PROTOCOLS

Consumer Product Safety Commission

Consumer Product Safety Act/Federal Hazardous Substances Act

Prior to passage of the Consumer Product Safety Act (U.S. Congress, 1972) household products were classified and tested for acute toxicity under FDA which administered the Federal Hazardous Substances Act (U.S. Congress, 1960). The role of administering the Federal Hazardous Substances Act (FHSA) now belongs to the Consumer Product Safety Commission (CPSC) under responsibilities of the Consumer Product Safety Act.

The primary purpose of the Consumer Product Safety Commission is the regulation of potentially hazardous household consumer products. The CPSC, however, has no statutory authority to mandate pre-market testing of consumer products by the manufacturer. The Commission's field offices collect products which are being sold in the market place and are suspected of noncompliance with labeling regulations and are therefore suspect hazardous substances. Consumer products are tested for hazard according to the test protocols listed in the FHSA (National Research Council, 1977).

The FHSA extended the government's power to ban substances for which risk could not be satisfactorily reduced by labeling. The FHSA is very specific in what is required for labeling purposes (Hadden, 1983). Hazardous substances are determined form toxicity tests and are categorized for labeling purposes as either highly toxic or toxic. Specific acute oral and inhalation test protocols are not detailed in the FHSA, however, methods of performing the following tests are specified in the regulations:

1) Acute dermal toxicity (single exposure)
2) Primary dermal irritation
3) Eye irritation.

If specifically prescribed criteria are met during the performance of any of these tests, labeling must take place as defined in the regulations. Under FHSA, data based on human experience, when available, is preferred over animal data and is given precedence over animal data when both exist and results differ.

Poison Prevention Packaging Act

The Poison Prevention Packaging Act (U.S. Congress, Poison Prevention Packaging Act, 1970) has as its primary concern the protection of children from poisonous household substances, such as pesticides, household cleaning agents, household fuels, and the like. The CPSC has the responsibility of administering this labeling act. No specific testing is detailed, but the regulations provide that labeling must comply with the requirements of the FHSA described above.

Lead-Based Paint Poison Prevention Act

The Lead-Based Paint Poison Prevention Act (U.S. Congress, 1973) also administered by the CPSC regulates the tolerance of lead in point and similar surface coating materials. This Act banned toys and other products used by children that bore lead-containing paint and the selling of furniture which bore lead-containing paint. The lack of specific testing requirements and use of the FHSA as a model are similar to that described for the Poison Prevention Packaging Act.

Department of Agriculture

No specific testing requirements were found at the agency or subagency level of the Department of Agriculture in this review. The following material summarizes the significant labeling acts administered by the USDA, which may lead to animal testing.

Egg Products Inspection Act

Under this Act (U.S. Congress, Egg Products Inspection Act, 1970) the Department of Agriculture must regulate all eggs and egg products used in interstate and foreign commerce. The quality of eggs must be inspected for the health and welfare of the consumer. The act provides for inspection of business premises, facilities, inventory, operation, records of egg handlers, and the like in order to assure that egg products are not adulterated and have been properly labeled and packaged.

Poultry Products Inspection Act

Under the Poultry Products Inspection Act (U.S. Congress, 1957) the Department of Agriculture is directed to regulate poultry and poultry products used in interstate and foreign commerce. The inspection provisions are similar to those cited above for the Egg Products Inspection Act.

Federal Meat Inspection Act

Under the Federal Meat Inspection Act (U.S. Congress, 1967), Congress mandates the Department of Agriculture to regulate all meat and meat food products used in interstate and foreign commerce. This act is targeted to protect the consumer by requiring the inspection of animals before slaughter, inspection of business facilities and slaughterhouses, and the inspection of all meat and meat food products. Any condemned meat or meat food product must be destroyed. The Federal Meat Inspection Act also specifies labeling and packaging requirements for meat products.

Department of Health and Human Services

FOOD AND DRUG ADMINISTRATION

Federal Food, Drug, and Cosmetic Act (FFDCA)

The Federal Food, Drug, and Cosmetic Act (U.S. Congress, 1938) covers the biological testing requirements for the general categories of food, food additives, drugs, devices (clinical chemistry and toxicology test systems, medical and dental diagnostic, prosthetic, surgical, therapeutic products, and hematology and pathology products, reagents and kits), and cosmetics. The overwhelming task of enforcing the law is accomplished by the FDA in the Department of Health and Human Services (DHHS). Under conditions where certain substances may be found in foodstuff and associated testing is required, the regulation occurs outside of the FFDCA (e.g., pesticides in agricultural commodities, FIFRA).

The Secretary of the DHHS is responsible for setting FFDCA regulations and oversees the Administrator of the FDA on such matters as: limiting quantities of deleterious materials in foodstuffs and determining whether a particular food additive, drug, or cosmetic is safe. Under the latter circumstances the DHHS must consider: (1) probable human consumption of, and/or topical exposure to, the substance, (2) cumulative effects of the substance on the diet and/or the skin, and (3) safety factors as recognized by scientific experts in animal experimentation.

The FDA chemical toxicology research is mainly associated with foods, drugs, and cosmetics. Testing programs include the following:

(1) acute, subchronic and chronic toxicity

(2) dermal and ocular toxicity

(3) development and reproductive toxicity

(4) immunotoxicity

(5) nutritional effects of toxicity

(6) disposition, metabolism, and pharmacokinetics of toxic substances

(7) genetic toxicity and genetic repair mechanisms.

These programs are actually guidelines for the determination of an adequate toxicity data base for each substance. The FDA handles each substance on a case-by-case basis. The exact testing regimen is determined by considering the type of substance (e.g., food additive), method of exposure, and the degree of intended use. These factors also influence the required duration of subchronic and chronic toxicity studies. Thus, the choices and protocols of the testing programs listed is different for each chemical substance.

Legal use of a food additive or color additive requires that industry submit a petition to FDA which shows whether or not it will be safe for its intended use. Specific toxicological testing protocols are not given by FDA, however. FDA specifies that cosmetics be free of any "poisonous and deleterious" substances, but does not required the premarket clearance of cosmetics.

The investigation of new drugs is now handled under the New Drug Regulations. Prior to clinical investigation of new drugs in humans a series of animal tests have been recommended by the FDA. The recommended testing format follows the normal phase sequence of clinical investigations. Again, this procedure is not strictly followed, as more or less extensive testing may be required depending upon considerations such as degree of intended use and segment of human population targeted (e.g., children). An outline of the recommended tests follows.

I. Drugs administered Orally or by Parenteral Injections

   A. Phase 1

      1. For clinical administration of a single dose or treatment over a period not to exceed a few days in any one individual:

a. Acute toxicity: 3-4 species (at least one a non-rodent) by the route to be employed clinically.

b. Subacute toxicity: two species (one a non-rodent) at 3 dosage levels for 2 weeks by the route to employed clinically. The following parameters should be considered:

(1) Body weights, food consumption, behavior

(2) Hematology (pre-testing and final)

 (a) Hemogram

 (b) Coagulation tests (non-rodent only)

(3) Clinical Chemistry (pre-testing and final)

 (a) Liver function: serum glutamic pyruvic transaminase: alkaline phosphatase (non-rodent only): others as indicated.

 (b) Kidney function: Urinalysis, BUN (non-rodent only).

 (c) Fasting blood sugar (both species).

(4) Ophthalmologic examination (pretesting and final).

(5) Gross and microscopic examination of major organs, and others as indicated.

c. Special study: Parenteral studies (i.e. tolerance, perivascular tolerance, compatibility with blood, etc.) where applicable.

2. For clinical tests requiring more than a few days treatment: more extensive animal studies as the case warrants: i.e., acute (as in 1.a. above) plus subacute (as for Phase 2).

B. Phase 2

1. For single clinical administration (e.g., diagnostic drugs, local anesthetics): no further studies beyond those designed to support Phase 1.
2. For clinical trial of 1 to 2 weeks duration: subacute toxicity in two species (at least one a non-rodent) at 3 dosage levels for 4 weeks, by the route to be employed clinically. Observations as for Phase 1; hematology and clinical chemistry prior to dosing and at 1, 2, and 4 weeks; ophthalmologic examination (pre-testing and final).
3. For 1 to 3 months clinical trial: no further studies beyond 2 above.
4. For 6 months to unlimited clinical trial: subacute toxicity in two species (one a non-rodent) at 3 dosage levels for 3 months by the route to be employed clinically. Observations as for Phase 1. Hematology and clinical chemistry prior to dosing, at 1, 2, and 4 weeks and monthly thereafter; ophthalmologic examination prior to dosing and at monthly intervals thereafter.

C. Phase 3

1. For single administration clinically: no further studies beyond Phase 1.

2 For 1 to 2 weeks clinical trial: subacute toxicity in two species (one a non-rodent) at 3 dosage levels for up to 3 months, depending upon the situation. Observations as for Phases 1 and 2.

3. For 1 to 3 months clinical trial: subacute toxicity tests as in 2 above, but of at least 3 months duration.

4. For 6 months to unlimited clinical trial: subacute toxicity in two species (one a non-rodent) at 3 dosage levels for at least 6 months. Depending on the circumstances, the duration of these experiments may have to be extended up to 2 years. Observations as for Phase 2.

II. General Anesthetics Administered by Inhalation. Phase 1, 2, and 3: Subacute studies in 4 species. Administer anesthetic for 3 hours daily on 5 consecutive days under conditions to be used clinically. Tests should include hematology, clinical chemistry, pathologic examinations, etc., as outlined above for subacute studies supporting Phase 1 trials of oral and parenteral drugs (A.1.b).

III. Drugs for Dermal Application

A. Phase 1

1. Acute toxicity: oral-two species (one a non-rodent) using formulation to be employed clinically.

2. Single administration clinically: dermal single exposure for 24 hours followed by observation period of at least 2 weeks. To support multiple clinical administrations, subacute studies are also needed (see Phase 2). Parameters monitored should include those mentioned under oral and parenteral drugs. Necropsy should include gross and microscopic inspection loci of application, as well as vicera.

B. Phase 2

1. Acute oral and dermal toxicity as for Phase 1.

2. Subacute toxicity: daily dermal application to intact and abraded skin for 3 weeks at approximately 1, 3, and 10 times the human dose on a body-weight basis. Observations as for Phase 1.
3. Guinea pig or human sensitization tests.

C. Phase 3

1. Short-term clinical trial. No further studies beyond Phase 2.
2. Unlimited trial. As for Phase 2, but subacute toxicity is extended to include a 3-6 month study on intact skin.

IV. Drugs for Topical Ophthalmic Administration

A. Phase 1

1. Acute toxicity: orally in 1 or 2 species (including the rat), using formulation to be employed clinically.
2. Irritation tests employing the rabbit eye
3. To support multiple clinical administrations: subacute studies are also needed (see Phase 2).

B. Phase 2. Subacute toxicity: three weeks daily applications (must be repeated several times daily if to be so used clinically). Parameters monitored should include those mentioned under oral and parenteral drugs. Special emphasis should be placed on ophthalmologic examination.

C. Phase 3. Depending upon clinical use, the subacute study should be extended, commensurate with the period of drug administration.

V. Drugs for Vaginal or Rectal Administration

A. Phase 1

1. Acute toxicity: orally in 1 or 2 species, including the rat

2. Local and systemic toxicity by single rectal or vaginal administration in the rat and a non-rodent
3. To support multiple clinical administrations, subacute studies are also needed (see Phases 2 and 3).

B. Phases 2 and 3. Subacute toxicity: daily administration by the appropriate route at several dosage levels in two species (one a non-rodent). Duration and number of applications are determined by proposed clinical use. Parameters monitored should include those mentioned under oral and parenteral drugs. Gross and microscopic inspection of the loci of application should be made in addition to the usual pathological examinations.

VI. Combinations of Two or More Drugs (where toxicity data are already available on each individually)

A. Phase 1. LD_{50} determination (whenever possible by the route recommended clinically) of the combination in the rat, compared to individual, concurrently determined LD_{50}'s for indication of possible drug interaction.

B. Phases 2 and 3. Subacute toxicity: 1 to 3-month studies in rats and dogs on the combination, by the route recommended clinically. Duration necessary depends upon the particular situation. Parameters monitored should include those mentioned under oral and parenteral drugs.

In addition to the programs for testing substances and providing regulatory decisions, the FDA is also mandated to conduct and sponsor research on Method Development. The purpose is to develop and validate procedures which are: (a) faster, (b) less expensive, and (c) more reliable in providing data more meaningful for extrapolation of toxic effects in animals to toxic effects in man. The FDA research program in method development and validation currently includes carcinogenesis, teratogenesis, mutagenesis, and toxicities involving reproductive, neurobehavioral, cardiovascular, dermal, endocrine, ocular and immune systems.

Fair Packaging & Labeling Act

The purpose of this Act (U.S. Congress, Fair Packaging & Labeling Act, 1966) is to protect consumers from economic victimization due to imperfect or purposely distorted information. It assures that labels of packaged consumer products inform consumers of quantity and contents.

NATIONAL TOXICOLOGY PROGRAM

The National Toxicology Program (NTP) was established in November, 1978, as a Department of Health and Human Services cooperative effort to strengthen activities related to the development and validation of new and improved toxicological testing methods. Toxicology research and testing activities within NTP include:

(1) Mutagenesis (cellular and genetic toxicology)
(2) Carcinogenesis
(3) Toxicologic characterization
(4) Chemical disposition
(5) Cutaneous toxicology
(6) Immunological toxicology
(7) Neurobehavioral toxicology

(8) Pulmonary toxicology

(9) Reproductive and developmental toxicology

NTP has, as one of its broad goals, the targeting of chemicals potentially toxic to man and the dissemination of this information to research and regulatory agencies (Hart et al., 1983).

NATIONAL INSTITUTE OF OCCUPATIONAL SAFETY AND HEALTH

Occupational Safety and Health Act (OSHA)

The OSHA (U.S. Congress, Occupational Safety and Health Act, 1970) mandates the National Institute of Occupational Safety and Health (NIOSH) to conduct Health Hazard Evaluations (HHE) of workplaces. On occasion HHE's reveal morbidity that cannot be explained on the basis of current toxicology knowledge. In response to these circumstances, NIOSH must design toxicological investigations for the purposes of elucidating mechanisms of toxicity that may account for the HHE findings. These areas are:

(1) mutagenesis screening (screen of occupational chemicals)

(2) reproductive toxicology

(3) carcinogenesis

(4) cutaneous toxicology

(5) neurobehavioral toxicology

(6) general toxicologic assays (acute, subchronic, chronic- for dose-response data)

All of these studies concern workplace substances. The main thrust of the research program is testing inhalation exposure. Additionally, nearly

25 percent of the NIOSH toxicology budget is devoted to method development, that is, the development of new and better assay systems.

Federal Mine Safety and Health Act

The Federal Mine Safety and Health Act (U.S. Congress, Federal Mine Safety and Health Act, 1977) mandates the health and safety standards for the protection of life and prevention of injuries in coal or other types of mines. The HEW Secretary must determine whether a toxic material or harmful physical agent which is found or used in a mine is potentially toxic at the concentration in which it is found or used. A base set of toxicological tests are used for these purposes, which can vary with the chemical substance in question.

Any mandatory health or safety standard shall prescribe the use of labels or other warnings as necessary to insure that miners know the hazards to which they are exposed, relevant symptoms, appropriate emergency treatment, proper conditions and precautions for safe use, protective equipment, and monitoring methodology.

Department of Housing and Urban Development

Lead-Based Paint Poison Prevention Act

The HUD, under the Lead Based Paint Poisoning and Prevention Act of 1973 (U.S. Congress, 1973), as amended 1976, is authorized to ensure that all federally subsidized homes meet the specifications of the act. No specific testing requirements have been specified.

Department of Transportation

MATERIALS TRANSPORTATION BUREAU

Hazardous Material Transportation Act

The Department of Transportation (DOT) requires that any material shipped via interstate commerce be properly labeled and contained in a manner reflecting the degree of hazard it presents. DOT generally regulates the transportation of toxic substances through the Hazardous Material Transportation Act (U.S. Congress, 19__). The DOT requires that acute toxicity studies be carried out on substances which are not already classified as either Class A Poisons ("extremely dangerous poisons") or Class C Poisons ("tear gases or irritating substances"), but which could be specified as Class B Poisons.

A Class B Poison label in transportation is required if the substance is known to be so highly toxic to man as to present a health hazard during transportation. A Class B label is also necessary if any one of the following conditions are met:

1) if, when administered orally at a single dose of 50mg or less per kilogram of body weight, death is produced within 48 hr. in half or more than half of a group of 10 or more white laboratory rats
2) if, when administered topically at a dose of 200mg or less for a period of 4 hours, death is produced within 48 in half or more than half of a group of 10 or more rabbits

3) if, when inhaled continuously for a period of one hour or less at a concentration of 2mg/L or less, death is produced within 48 hours.

COAST GUARD

Port and Waterways Safety Act

Given the responsibility of protecting U.S. waterways (U.S. Congress, 19__), the U.S. Coast Guard, within the Department of Transportation, requires the following mammalian acute toxicity data before a chemical may be imported into the United States: 1) acute oral and dermal toxicity, 2) dermal and eye irritation studies.

FEDERAL RAILWAY ADMINISTRATION

Federal Railway Safety Act

The Federal Railway Safety Act (U.S. Congress, Federal Railway Safety Act, 1970) regulates the transportation of toxic substances on rail. No testing requirements have been specified.

Environmental Protection Agency

Clean Water Act

The Clean Water Act (U.S. Congress, 19__) calls for deriving National Water Quality Criteria. The complete data set is developed by conducting a series of acute and long-term bioassays using organisms from at least eight different families. The required acute tests pertinent to this study are as follows:

1. A salmonid
2. Another family belonging to the class Osteichthyes (bony fishes)
3. Another representative of the phylum Chordata

The following long-term tests are also required:

1. Chronic tests with one fish
2. A bioconcentration test conducted with one aquatic species.

In October, 1982, EPA published a Draft Water Quality Standards Handbook (U.S. Environmental Protection Agency, Draft..., 1982). This document was published as proposed rules known as "Water Quality Standards Regulation" under Section 303 of the Clean Water Act. This handbook provides guidance for developing site-specific water quality criteria which reflect local environmental conditions. This process considers the species that are resident at particular sites and the water quality characteristics of that site. There are four procedures in these Guidelines for developing site-specific criteria: the recalculation procedure, indicator species procedure, resident species procedure, and heavy metal speciation procedure. With the exception of the heavy metal recalculation procedure, the remaining above-listed procedures each require fish testing to develop site-specific criteria.

Recalculation Procedure. This procedure is designed to compensate for sensitivity differences between resident species data, which may be available on a national level. Bioassay testing is necessary if, after eliminating data for inappropriate species, the minimum data set requirements are not met. Under these conditions, additional acute toxicity testing in laboratory water with resident fish is required to calculate a final acute value for that site. In addition, for non-lipid soluable chemicals whose residue value is controlled by an FDA action level, a bioconcentration test with an edible resident aquatic species is required to determine the site-specific residue value.

Indicator Species Procedure. This procedure is based on the premise that water quality at a particular site may influence the toxicity and/or bioavailability of a chemical. To determine a site-specific maximum instantaneous concentration, acute toxicity tests with a sensitive fish species must be conducted in both laboratory dilution water and dilution water from the site. To determine a site-specific chronic value, a State may conduct up to three chronic tests with sensitive fish in laboratory dilution water and up to three chronic tests in dilution water from the site.

Resident Species Procedure. The Resident Species Procedure is designed to compensate for both species sensitivity and differences in water quality. To utilize this procedure a State must produce a complete acute toxicity data set using resident species in water from the site. The actual number of tests that may be required is variable, depending upon the number of species present. The resident species chronic value is determined by conducting up to three chronic tests with sensitive fish in the laboratory with dilution water from the site. More than one set of acute and chronic tests may be required, depending upon seasonal changes in water quality. Such changes in water chemistry and biota are known to alter responses of such fish to toxic substances.

Clean Air Act

The Clean Air Act (U.S. Congress, Clean Air Act, 1970) has as its aim the protection and enhancement of the quality of the nation's air resources thereby promoting the public health and welfare and the productive capacity of its population. As required by Section 110 of the 1970 Clean Air Act, each state must have a plan which provides for implementation, maintenance, and enforcement of the National Ambient Air Quality Standards in that state. Section 112 authorized EPA to establish a list of hazardous air pollutants and to promulgate emission or design standards for both new and existing sources which emit these pollutants. Hazardous air pollutants are ones for which no ambient air quality standard has been set but which may cause an increase in mortality or serious illness. Carcinogens meet this definition and Section 112 is a likely method for regulating their emission.

For registration of any fuel or a fuel additive, the EPA Administrator may require the fuel/fuel additive manufacturer to conduct tests to determine any potential public health effects of a fuel/fuel additive on both a short-term and long-term basis. Acute and chronic studies are performed. In addition, investigations are conducted for assessment of immunotoxicity, carcinogenesis, teratogenesis, and mutagenesis (including, but not limited to, carcinogenic, teratogenic, or mutagenic effects).

Resource Conservation and Recovery Act

The Resource Conservation and Recovery Act (U.S. Congress, Resource Conservation and Recovery Act, 1976) was designed to protect human health and the environment from exposure to hazardous waste. Section 3001 of Subtitle C directs EPA to characterize and list those hazardous wastes that are subject to regulation. Details of the hazards are obtained from existing literature; when information does not exist for certain wastes, it is the responsibility of EPA to develop it. The Resource Conservation and Recovery Act (RCRA) does not require generators of hazardous wastes to test the toxicity of their wastes.

Under RCRA animal testing is required to determine if there are any hazards associated with a particular waste. Since RCRA only deals with solid wastes, the predominant ecological problems arise from the leaching of wastes from disposal sites. At present, EPA has not promulgated specific tests for the toxicological characterization of wastes, but is now in the process of selecting and validating test procedures for this task. The RCRA characterization procedure includes testing to determine both acute and chronic toxic effects on aquatic animals.

In terms of aquatic tests, static acute bioassays with fish are required. The primary species used for these bioassays are fathead minnows. Partial or full life cycle bioassays with fathead minnows are also required under RCRA. In addition, long-term effects are also investigated by conducting a fish bioaccumulation test.

In addition to environmental testing requirements, RCRA requires that leachates of wastes be characterized in terms of their potential hazard to man. One test from each of the following groups is required:

- Group I -point mutation in bacteria;
- Group II - cultured mammalian somatic cells or fungal microorganisms;
- Group III - DNA repair in bacteria; unscheduled DNA synthesis in human diploid cells, sister-chromatid exchange in mammalian cells, or mitotic recombination and/or gene conversion in yeast.

Safe Drinking Water Act

The Safe Drinking Water Act (U.S. Congress, 1974) provides for the determination of the maximum allowable contaminant levels of various chemical substances in public water systems which are used for human consumption.

Toxic Substances Control Act

Section 4 of the Toxic Substances Control Act (U.S. Congress, Toxic Substances Control Act, 1976) governs test rules and guidelines to be carried out by the EPA administrator. The Toxic Substances Control Act (TSCA) recognizes the need for testing of diverse life forms at various levels of biological organization. These levels are subcellular, cellular, organismal, population, community, and ecosystem. Presently, emphasis is placed upon single species test populations of vertebrates. The reason for this emphasis is that the EPA considers such single species testing to be state of the art of animal testing methodology (U.S. Environmental Effects Protection Agency, Health..., 1982).

Under TSCA, an Interagency Testing Committee (ITC) is authorized to designate up to 50 chemical substances or groups for which the EPA Administrator must within 12 months either initiate rulemaking requiring their testing or publish reasons for not taking such action. This has been termed the Jellinek Schedule. The ITC is comprised of representatives from EPA, OSHA, NIOSH, CEQ, NIEHS, NCI, NSF, and DOC. Table 1 lists what the ITC must consider under TSCA when in the process of designating chemical substances to the priority list. Priority attention is given by the ITC to substances or mixtures which are known or suspected carcinogens, mutagens, or teratogens.

When a "new"[1] chemical substance is under consideration for production, a premanufacture notification (PMN) must be submitted to the EPA. This must be submitted at least 90 days before the manufacturer commences with production. The EPA recommended starting point for the premanufacture testing program, to be submitted with the PMN is listed in Table 2.

These data are collected, packaged, and sent to the EPA along with PMN forms. The EPA can then make a decision on any further tests required. These tests would be teratogenic, carcinogenic, subchronic, and/or chronic as needed. The EPA recommends that the data in Table 2 be obtained from methods publications of the Organization for Economic Cooperation and Development (OECD), the IRLG (Interagency Regulatory Liaison Group), or by EPA test standards of Section 4 of TSCA, Section 3 of FIFRA, or other approved EPA methods.

Health effects test requirements issued by EPA under TSCA fall into five categories. These categories are selected by the ITC for testing priority list substances:

(1) General Toxicity Testing (Acute, Subchronic, Chronic)

(2) Specific Organ/Tissue Toxicity (Includes Teratogenicity)

TABLE 1.

(1) Production volume expected

(2) Environmental release expected

(3) Potential occupational exposure of humans

(4) Potential non-occupational exposure of humans

(5) Similarity of the substance or mixture in question to others known to present unreasonable risk of injury to health or the environment

(6) Extent of data on the substance's health effects and extent to which additional testing of the substances will produce data for which effects can be predicted

(7) Reasonably foreseeable availability of facilities and personnel for performing the test being recommended

TABLE 2.

(1) Abundance of Physical/Chemical Data

(2) Acute Toxicity Data

Acute oral toxicity
Acute dermal toxicity
Acute inhalation toxicity
Dermal irritation
Dermal sensitization
Eye irritation (for chemicals showing no dermal irritation)

(3) Repeated Dose Toxicity Data
OECD - 14-28 days, repeated dose test(s) using probable route(s) of human exposure
TSCA - minimum 90 days, repeated dose test(s) using probable route(s) of human exposure

(4) Mutagenicity Data (Screening Tests)

Gene (point) mutation, chromosome aberrations

(5) Ecotoxicity Data

Acute toxicity (fish 96 hr LC_{50})

(6) Degradation/Accumulation Data

Ready degradability
Bioaccumulation (uptake from medium)

---

[1] "new" means not in the TSCA Inventory of Chemical Substances.

(3) Mutagenicity (Gene Mutations, Chromosomal Effects, DNA Effects)

(4) Neurotoxicity

(5) Special Studies (Metabolism)

The purpose and importance of the above tests is discussed later in this report. A summary of test protocols follows in Table 3.

Federal Insecticide Fungicide and Rodenticide Act

This Act is designed to protect human health and the environment, including fish and wildlife, from the misuse of pesticides, herbicides, fungicides, and rodenticides (U.S. Congress, 1947).

Human Health Testing. Human health testing requirements under FIFRA are quite stringent. The specific requirements for tests are dependent upon the individual test compound. However, the tests, given in Table 4 are generally required as a minimum for meeting FIFRA requirements, and additional testing may be necessary if the test compound has the potential to produce effects not covered by these tests. For most of the tests, mice or rats are the preferred test animals, however, in certain instances, rabbits, guinea pigs, dogs, chickens, or primates are also acceptable. In many substances, more than one species is required for testing which depends upon the characteristics of the test compound.

Fish and Wildlife Testing. The fish and wildlife testing requirements for the registration of a pesticide or herbicide are broken down into avian, wild mammal, and aquatic organism tests. In the case of fish and wildlife a tiered approach is required, which allows the results of each tier

TABLE 3. SUMMARY OF TEST PROTOCOLS GENERALLY REQUIRED UNDER TSCA AND FIFRA

Type of Test	Species Used	# Animals per Dose Level	# Dose Levels	Duration of Test
Acute Dermal Toxicity	Rat, Rabbit, Guinea Pig (Rabbit preferred)	10 (5F, 5M)	At least 3	24 hr. exposure; 14 day observation period
Acute Inhalation Toxicity	Rat	10 (5F, 5M)	At least 3	4 hr. exposure; 14 day observation period
Acute Oral Toxicity	Rat	10 (5F, 5M)	At least 3	Single dose in 24 hr. period; 14 day observation
Repeated Dose Dermal Toxicity*	Rabbit			21-28 days
Repeated Dose Inhalation Toxicity*	Rodent			14 or 28 days
Repeated Dose Oral Toxicity*	Rodent or chicken			14 or 28 days
Subchronic Dermal Toxicity	Rat, albino rabbit, guinea pig (rabbit preferred)	20 (10F, 10M)	3 + control	90 days
Subchronic Inhalation Toxicity	Rat	20 (10F, 10M)	3 + control	90 days
Subchronic Oral Toxicity	Rat, Dog	20 (10F, 10M) - Rodent 8 (4F, 4M) - Non-Rodent	3 + control	90 days
Chronic Toxicity	2 mammalian species one rodent (rat) and one non-rodent (dog)	40 (20F, 20M) - Rodent 8 (4F, 4M) - Non-Rodent	3 + control	12 months
Dermal Sensitization**	Adult guinea pig	Dependent on test method chosen	Dependent on test method	Dependent on test method
Primary Dermal Irritation	Albino rabbit	6	1	4 hrs. exposure; at least 72 hrs. observation up to a maximum of 14 days
Primary Eye Irritation	Albino rabbit	6	1 (one treated eye per animal one untreated eye per animal as control	72 hrs. after application
Reproduction/Fertility Effects	Rat	20 M and sufficient # of F to yield at least 20 pregnant F at or near term	3 + control	28 weeks from dosing of parental animals
Teratogenicity	2 mammalian species Rat and rabbit preferred	20 rats, 12 rabbits	3 + control	
Carcinogenicity	Rat or mouse	100 (50F, 50M)	3 + control	24 mos. (rat), 18 mos. (mouse)
Combined Chronic Toxicity/Carcinogencity	Rodent (rat or mouse)	100 (50F, 50M)	3 + control	24 mos. (rat), 18 mos. (mouse)
Mutagenicity:				
A. Mouse Specific Locus Test	Mouse	Not specified (individual choice)	1	Not specified
B. *In Vivo* Mammalian Bone Marrow Cytogenetics Tests: Chromosomal Analysis	Rats; Mice; Chinese, Syrian, or American Hamsters	10 (5F, 5M)	3(also negative & positive controls)	Not specified

TABLE 3. (Continued)

Type of Test	Species Used	# Animals per Dose Level	# Dose Levels	Duration of Test
C. *In Vivo* Mammalian Bone Marrow *Cytogenetics* Tests: Micronucleus Assay	Mouse	10 (5F, 5M)	3(also negative & positive controls)	Bone marrow samples taken 3 times, not earlier than 12 hr. after treatment and not extending beyond 72 hrs.
D. Gene mutations of Somatic cells in culture	Mouse lymphoma cells Chinese hamster cells	NA	Multiple concentrations of test substance based upon cytotoxicity	Dependent upon cell line employed in test
E. Rodent Dominant Lethal Assay	Rat or mouse	# of treated males based on # of females to which male is mated	At least 3 (positive and negative controls)	Not specified
F. Rodent Heritable Translocation Assay	Mouse	At least 300 progeny per dose	At least 2 + control	Not specified
G. *In Vitro* Sister Chromatid *Exchange* Assay	Rodent	NA	Multiple concentrations of test substance over range adequate to define response (positive and negative controls)	Not specified
H. *In Vivo* Sister Chromatid *Exchange* Assay	Mouse, hamster, rat	NA	Multiple concentrations of test substance over range adequate to define response (positive or negative controls)	Not specified
I. Unscheduled DNA Synthesis in Mammalian Cells in Culture	Rat	Six independent cultures	Multiple concentrations of test substance, based upon cytotoxicity, and over a range adequate to define response	Not specified
Neurotoxicity:				
A. Neuropathology	Rat	6-10 (both sexes)	3 + control	Will be specified in test rule but generally is 90 days
B. Peripheral Nerve Function	Laboratory rodent	20 (either sex)	3 + control (also positive control)	Will be specified in test rule but generally is 90 days
C. Motor Activity	Laboratory rat or mouse	10	3 + control (also positive control)	Will be specified in test rule
D. Acute Delayed Neurotoxicity of Organophosphorus Substances	Domestic Hen	10 (sufficient # so that at least 6 survive the observation period)	3 + control	At least 21 days with re-dosing & observation for additional 21 days if no effects are seen
E. Subchronic Delayed Neurotoxicity of Organophosphorus Substances	Domestic Hen	10	3 + control	90 days
Metabolism	Rat	8 (4F, 4M)	2	7 day observation period after radioactive labeled dose or until 95% of dose has been excreted

* Section 3 under FIFRA.

** Any of the following test methods may be used: Freund's Complete Adjuvant Test, Guinea Pig Maximization, Split Adjuvant, Buehler, Open Epicutaneous Test, Mauer Optimizatoin, Footpad Technique in Guinea Pig.

From: U.S. Environmental Protection Agency, Health ..., 1982.

TABLE 4. FIFRA HUMAN HEALTH TESTING REQUIREMENTS

---

1. Acute oral toxicity - rat and mouse
2. Acute dermal toxicity - rabbit
3. Acute inhalation toxicity - rat
4. Acute dermal irritation/corrosion - rat
5. Acute eye irritation/corrosion - rabbit
6. Dermal sensitization - rabbit or guinea pig
7. Repeated dose 28-day or 14-day oral toxicity - rodent or chicken
8. Subchronic 90-day oral toxicity - rodent
9. Subchronic 90-day oral toxicity - non rodent
10. Repeated dose 21-28 day dermal toxicity - rabbit
11. Subchronic 90-day inhalation study - rodent
12. Repeated dose 14 or 28 day inhalation study - rodent
13. Subchronic 90-day inhalation study - rodent
14. Teratogenicity - rodent
15. Carcinogenicity - two years with rat or mouse
16. Chronic Toxicity - three generations with rat or mouse
17. Combined chronic toxicity/carcinogenicity - non rodent
18. Metabolism and residue accumulation - rat and larger mammals

---

of tests to be evaluated to determine if further testing is required. EPA expects the tests in the lower tiers will be required infrequently. The tests progress from simple laboratory exposures to applied field tests. Figures 1-3 provide the general information regarding this tiered approach.

Tier 2 calls for "Special Tests," which may be required under unusual circumstances and will be tailored specifically to that compound. Some examples of "Special Test" include: certain chemical properties, avian acute dermal LD_{50}, rabbit dermal LD_{50}, eye irritation, acute inhalation LC_{50} and fish or bird cholinesterase tests.

The requirements for each type of test for birds, wild mammals and aquatic organisms are presented below and will accompany Figures 1-3:

Avian Testing (See Figure 1).

Tier 1. Single Dose Oral LD_{50} Study - This test is required for all end-use formulated and manufacturing-use products. Avian Dietary LC_{50} Study - Required for all end-use and manufacturing use products.

Tier 2. Avian Reproduction Study - This study is required for all end-use formulated and manufacturing-use products if: (a) birds may be continuously or repeatedly exposed, (b) the product or its metabolites are stable in the environment or accumulated in plant or animal tissues or, (c) if mammalian reproduction is adversely affected by the compound. An avian reproduction study is generally conducted using either bobwhite quail or mallard ducks and requires approximately 22 weeks for completion. This test requires a minimum of 3 dose levels and a control. At least 8-12 replicate pens containing both male and female birds must be tested at each dose level.

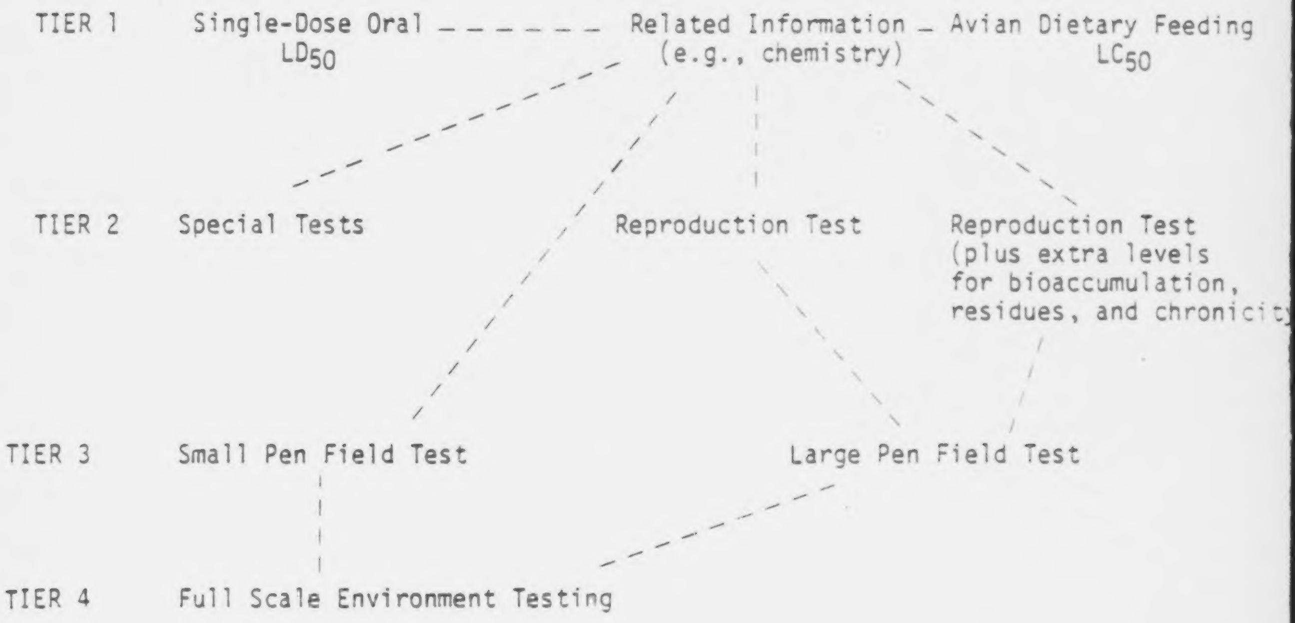
TIER 1
Single-Dose Oral LD_{50}
Related Information (e.g., chemistry)
Avian Dietary Feeding LC_{50}
TIER 2
Special Tests
Reproduction Test
Reproduction Test (plus extra levels for bioaccumulation, residues, and chronicit
TIER 3
Small Pen Field Test
Large Pen Field Test
TIER 4
Full Scale Environment Testing

FIGURE 1. AVIAN TESTING FLOW CHART

Tier 3. Large or Small Pen Field Testing - The need for these kind of tests is evaluated on a case-by-case basis. Specifically, these tests will be required if the above tests indicate that there may be adverse effects to birds that may be exposed to the pesticide.

Tier 4. Full-Scale Environmental Testing - The need for these tests are evaluated on a case-by-case basis. Generally, if the results of the preceding laboratory and field experiments have indicated negative effects at normal use rates, or evidence of long-term adverse effects, then this test will be required.

Wild Mammal Testing (See Figure 2).

Tier 1. Related Information - Information is required for all end-use formulated and manufacturing-use products.

Tier 2. Wild Mammal Toxicity Tests - These tests are required on a case-by-case basis. Hazard data for humans and domestic animals are normally adequate. Exceptions may be when there is variation in the sensitivity of mammalian species, when wild mammals may be heavily exposed to the product or when the product has bactericidal properties which may interfere with digestion of wild rumenants when consumed with foliage. Reproduction Tests - This is covered under tests required for human safety.

Tier 3. Large or Small Pen Field Testing - The need for these kinds of tests is evaluated on a case-by-case basis. These tests will be required if the above tests indicate that there may be adverse effects to wild mammals that may be exposed to the pesticide.

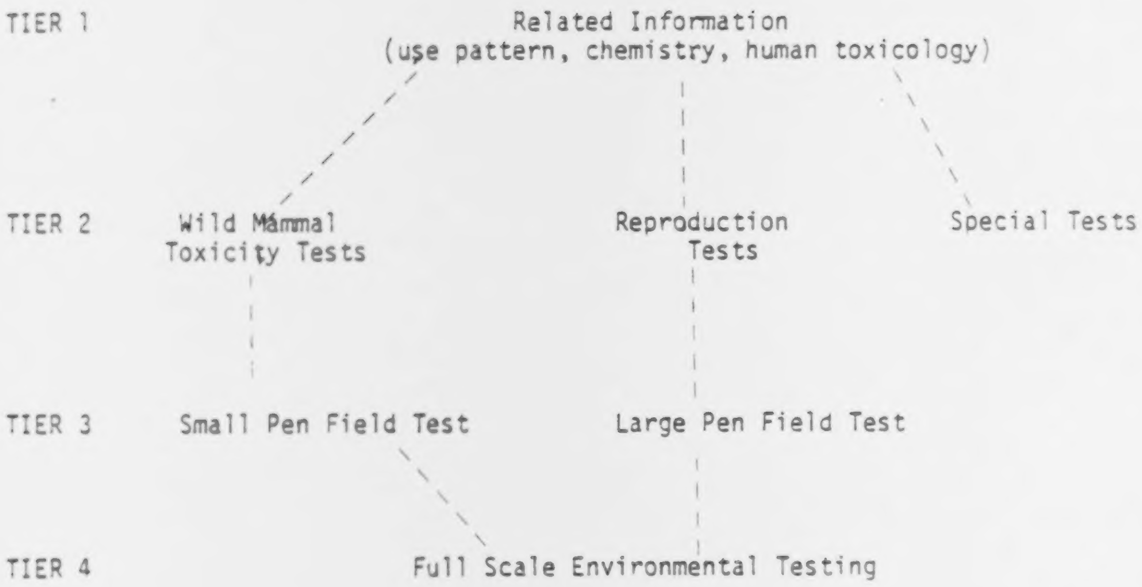
TIER 1
Related Information
(use pattern, chemistry, human toxicology)
TIER 2
Wild Mammal Toxicity Tests
Reproduction Tests
Special Tests
TIER 3
Small Pen Field Test
Large Pen Field Test
TIER 4
Full Scale Environmental Testing

FIGURE 2. WILD MAMMAL TESTING FLOW CHART

Tier 4. Full-Scale Environmental Testing - The need for these tests are evaluated on a case-by-case basis. If the results of the preceding laboratory and field experiments have indicated negative effects at normal use rates, or evidence of long-term adverse effects, then this test will be required.

Aquatic Organism Testing.

Tier 1. Acute Toxicity Test with Freshwater Fish - This short-term test is required for all end-use formulated and manufacturing-use products. Related Information - Information is required for all end-use formulated and manufacturing-use products.

Tier 2. Embryo-larval Tests - Long-term (30 d) tests are required for all end-use formulated and manufacturing-use products intended for use in water or that will be transported to water. The embryo-larval test is also required if (a) it will be continuously present in water, (b) the LC_{50} is 1 mg/L, (c) if the estimated concentration in water is 0.01 of the LC_{50} or (d) if the pesticide has the potential to bioaccumulate.

Acute toxicity of estuarine and marine organisms - Short-term tests of salt water organisims are required for all end-use formulated and manufacturing-use products which are intended for direct application to the estuarine or marine environment or they it will enter this environment in significant concentrations because of their use or mobility.

Aquatic Organisms Accumulation Test - Biouptake tests are required for all end-use formulated and manufacturing use products which are likely to

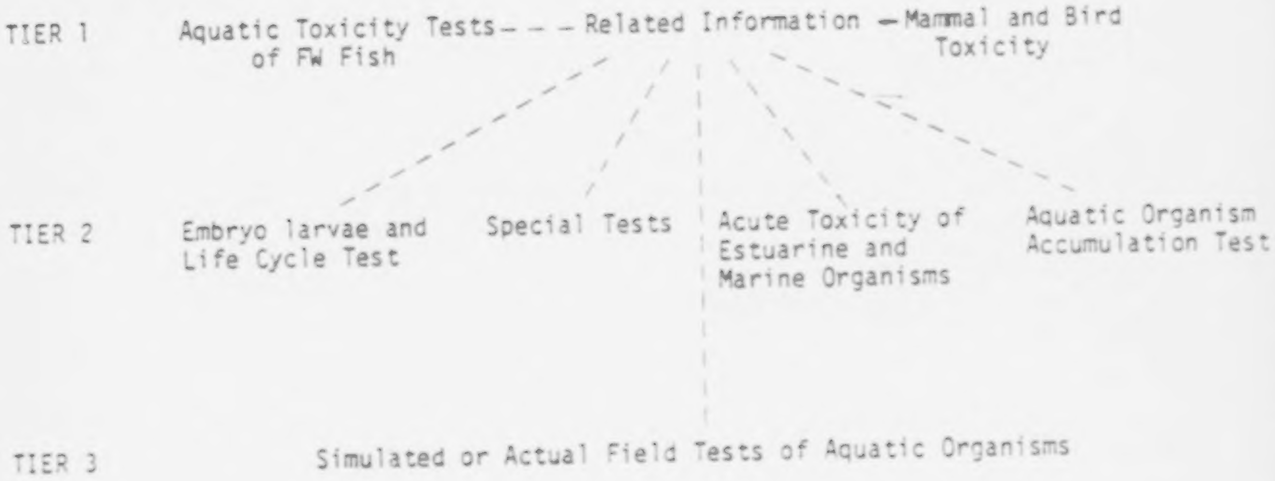
TIER 1
Aquatic Toxicity Tests of FW Fish
Related Information
Mammal and Bird Toxicity
TIER 2
Embryo larvae and Life Cycle Test
Special Tests
Acute Toxicity of Estuarine and Marine Organisms
Aquatic Organism Accumulation Test
TIER 3
Simulated or Actual Field Tests of Aquatic Organisms

FIGURE 3. AQUATIC ORGANISM TESTING FLOW CHART

result in residues in the aquatic environment which may accumulate in aquatic organisms to toxic levels. This is evaluated case-by-case but is usually required if: (a) the active ingredient has a water solubility of less than 0.5 mg/ and an octanol/water partition coefficient greater than 1000, (b) it is persistent in water, and (c) there is evidence that the product accumulates in mammals or birds.

Fish Life Cycle Tests - Full life cycle tests are required for all end-use formulated and manufacturing-use products if: (a) the estimated environmental concentration is one-tenth of the "no observed effect" level in the fish embryo-larval test, (b) if studies of other organisms indicate the reproductive physiology of fish may be affected.

Tier 3. Simulated or Actual Field Tests of Aquatic Organism - The need for these tests is evaluated on a case-by-case basis. Short-term field tests are required if the product is likely to cause adverse acute effects based upon available date, use pattern and exposure rate. Long-term tests are required if the available data suggest adverse long-term, cumulative or reproductive effects may result from the intended use.

Summary of Avian, Wild Mammal and Fish Testing Requirements. In practice, what this all means is that producers of pesticides, rodenticides, fungicides and herbicides typically submit results from tests for human health hazard and data from the following battery of tests, and wait for a response from EPA:

1. Bird single-dose oral LD_{50} with either quail, pheasant or mallard duck

2 & 3. Two avian dietary feeding LC_{50} tests, one with pheasant or quail and one with mallard duck.

4 & 5. Two acute toxicity tests, one with rainbow trout and one with bluegill.

Specific protocol requirements under FIFRA for the above tests are outlined in Table 5 (U.S. Environmental Protection Agency, Environmental ..., 1982). Avian, wild mammal, and aquatic organism testing requirements under FIFRA are summarized in Table 6. The results of these tests allow the EPA to set the maximum acceptable concentration of each test substance allowable in the environment.

SUMMARY

In summary, many of the statutes lack specific testing details. In many cases data submission requirements and specific testing protocols are left solely to the discretion of the regulators of the various agencies.

FOREIGN REGULATORY REQUIREMENTS

DATA SUBMISSION REQUIREMENTS

European Economic Community (EEC)

The 6th Amendment to council Directive 67/548/EEC (European Economic Community, 1979) establishes the regulatory framework within the European Economic Community for the regulation of toxic chemicals. Specific pre-market testing requirements for new chemicals are established by the 6th Amendment but existing chemicals are not included. If a manufacturer wishes to market a

TABLE 5. SUMMARY OF TEST PROTOCOLS GENERALLY REQUIRED FOR BIRDS AND FISH UNDER FIFRA

Type of Test	Species Used	#Animals per Dose Level	#Dose Levels	Duration of Test
Avian Single Dose LD_{50}	Bobwhite quail, pheasant, Mallard duck	10	((5)? + control)	Single dose in a 24 hr. period, 14 or 21 day observ tion period
Avian Dietary Feeding LC_{50}	Bobwhite quail, pheasant, Mallard duck	10 per dose level 30 per concurrent control 20 per negative control if applicable 20 per carrier control if applicable	5 + control	5 days exposure to test substance plus 3 days of additional observation
Fish Acute Toxicity Test	Rainbow trout Bluegill Fathead minnow	20 divided into 2 or more replicates	5 + control	96 hrs.
Fish Bioaccumulation Test	Fathead Minnow	Dependent upon expected length of tesing, sample size, and the number of additional specialized analyses to be performed at termination	1 constant sublethal concentration	Maximum 20 day uptake period; depuration period sho continue until 95% accumulated residues have been eliminated

From: U.S. Environmental Protection Agency, Environmental ..., 1982.

TABLE 6. GENERAL TESTING REQUIREMENTS UNDER FIFRA

Test Type	Required for All New Products	Evaluation on a Case-by-Case Basis
Avian		
Single Dose Oral LD_{50}	x	
Related Information	x	
Dietary Feeding LD_{50}	x	
Special Tests		x
Reproduction Test		x
Bioaccumulation Test		x
Large & Small Pen Field Testing		x
Full-Scale Environmental Test		x
Wild Mammals		
Related Information	x	
Toxicity Tests		x
Special Tests		x
Reproduction Tests		x
Small & Large Pen Field Testing		x
Full-scale Environmental Test		x
Aquatic Organisms (Fish)		
Acute Toxicity	x	
Related Information	x	
Special Tests		x
Embryo Larval Tests		x
Life-cycle Tests		x
Bioaccumulation Test		x
Simulated or Actual Field Tests		x

new chemical substance in a volume of 1-10 metric tons per year or a total market volume of less than 50 metric tons, he must submit a technical dossier containing "base set" information on the new substance to the authorities in one of the ten member EEC countries 45 days prior to marketing. The dossier serves to supply information which authorities in the member countries need in order to evaluate the immediate or delayed risks to man and the environment (Hushon et al., 1983). The following "base set" toxicological studies are required:

1. Acute oral toxicity (rat)
2. Acute inhalation toxicity (rat)
3. Acute percutaneous absorption (rat)
4. Skin irritation (albino rabbit)
5. Eye irritation (albino rabbit)
6. Skin sensitization (guinea pig)
7. Subchronic oral toxicity (rat)
8. Subchronic inhalation toxicity (rat)
9. Mutagenicity Studies
10. Ecotoxicological Studies (fish)
11. Degradation Studies

When marketing of a new chemical reaches 10 metric tons per year or a total of 50 metric tons, Level 1 has been reached according to the 6th Amendment. After careful consideration of the current knowledge of the substance, its present and potential uses, and the results of the "base set" studies the following additional studies may be required:

1. One generation fertility study
2. Teratology study
3. Additional subchronic and/or chronic toxicity studies which can also include special studies (90 day exposures up to 2 years). Effects which would warrant such a study include (a) serious or irreversible lesions (b) a very low or absence of a "no effect" level (c) a clear chemical structure relationship between the substance being tested and other substances which have been proven dangerous
4. Additional mutagenesis studies (including carcinogenesis screening)

If results of the mutagenesis tests are negative, a test to verify mutagenesis and to verify carcinogenesis is required. If the mutagenesis verification test is also negative, no further mutagenesis testing is required at Level 1. If the test results are positive, however, further tests must be carried out.

Level 2 in the 6th Amendment testing regime is reached when the quantity of a marketed substance reaches a level of 100 metric tons per year or a total of 500 metric tons. Level 2 testing includes the following:

1. Chronic toxicity study
2. Carcinogenicity study
3. Three generation fertility study if an effect on fertility was established at Level 1
4. Teratology study (non-rodent) to verify results of the Level 1 teratology study
5. Acute and subchronic toxicity studies on a second species if Level 1 test results indicate a need (Results of biotransformation and pharmacokinetics studies can also indicate that additional acute and subchronic testing is required)

6. Ecological toxicity assessment including prolonged toxicity studies with fish (also including reproductive studies)

An example of data required for the testing of new drugs is given in Table 7.

Non-EEC Countries

Japan

The data submission requirements in Japan are similar in nature to those of the EEC. However, the exact information obtained with regard to species, dose and duration may differ (Table 8).

METHODOLOGIES AND PROTOCOLS

World Health Organization

The World Health Organization (WHO) was established in 1948 as one of the specialized agencies closely allied with the United Nations. WHO deals with problems which can only be solved through the cooperation of all or several countries working in concert. Through WHO, the public health and medical professions of over 150 countries exchange knowledge and work together in an effort to achieve the highest possible level of health throughout the world. WHO's major thrust areas include the testing and evaluating of food additives and pesticides.

All intentional food additives should be subjected to individual toxicological investigation (Table 9). Adequate specifications for purity and identity of the food additive should be available before toxicological work is undertaken. WHO recommends that toxicological evaluation studies be conducted

TABLE 7. EEC PROTOCOLS FOR TESTING NEW DRUGS

Type of Test	Species Used	#Animals per Dose Level	#Dose Levels	Duration of Test
Single dose acute toxicity	2 of known strain (rodent and non-rodent)	Not specified	Rodents: As many as necessary to calculate LD_{50} with 95% confidence limits Non-rodents: As many as necessary to assess spectrum of toxicity & estimate lethality	14 days or as long as signs of toxicity persist
Repeated dose toxicity	At least 2, one a non-rodent, to be chosen on the basis of their similarity to man	Not specified	3	According to proposed duration human treatment: In man: / In a 1 or repeated doses in 1 day / 2 Up to 7 days / 4 Up to 30 days / 3 mo
Reproduction study	2 for embryo toxicity studies, at least for fertility & perinatal studies one of the species (if possible) that is used in long-term toxicity studies, third species if conflicting results obtained	Embryo toxicity: 20 pregnant Females (Rodent) 12 pregnant Females (Non-rodent) Fertility: 24 Females/24 Males Perinatal: 12 pregnant Females	3	Not specified
Carcinogenicity	At least 2, preferably those with metabolism of drug similar to man	Not specified	3	Not specified

From: Alder et al., 1981.

TABLE 8. JAPAN PROTOCOLS FOR TESTING NEW DRUGS

Type of Test	Species Used	#Animals per Dose Level	#Dose Levels	Duration of Test
Acute toxicity	At least 2 (not specified)	not specified	not specified	at least 72 hrs.
Subchronic toxicity	At least 1 (not specified)	not specified	Dose has to be lethal to animals in a month 4 dose levels, including control	One month or more 3 months or more
Fetal tests				
A) Administration before & at the start of pregnancy	At least 1 (rat or mouse)	20 M, 20 F	Several, including maximum dose that adult animal can tolerate	Not specified
B) Administration tests during the organogenic period of fetus	At least 1 rodent (rat or mouse) and one non-rodent (rabbit)	30 rats or mice 10 rabbits	Several, including maximum dose that adult female can tolerate	Not specified
C) Administration tests during last part of pregnancy & nursing period	At least one (mouse, rat, or rabbit)	Rodents: 20 Non-rodents: 10	Several, including maximum dose that adult animal can tolerate	Not specified
Carcinogenicity	OFFICIAL GUIDELINES (IF ANY) COULD NOT BE OBTAINED			

From: Alder et al., 1981.

TABLE 9. RECOMMENDED PROTOCOLS FOR TESTING FOOD ADDITIVIES

Type of Test	Species Used	#Animals per Dose Level	#Dose Levels	Duration of Test
Acute Oral Toxicity	At least 3, one should be non-rodent	Not specified	Note specified	2-4 weeks
Subacute toxicity	Rodent (rat) and non-rodent (dog)	10-20 of each sex (rodent) 2 of each sex (non-rodent)	Sufficient # to ensure that at least one level has no effect and that doses are provided which show definite toxic effects. A control should be used.	90 days-rodent 12 months-non-rodent
Chronic toxicity	Rat	25 of each sex	2 + control	18-24 months

From: World Health Organization, 1957; Ibid, 1958; Ibid, 1967.

using the species most similar to man with regard to its metabolic, biochemical, and toxicological characteristics in relation to the substance under test. For example, if the metabolism of the additive in man is found to be significantly different from that in rodents (most widely accepted test species), investigation with other species may be necessary.

Flexibility of approach is essential in deciding the duration of animal toxicity tests for the evaluation of an individual food additive. All test reports should include the following: (1) species (2) sex (3) diet (4) route of administration (5) duration of test.

According to WHO, an assessment of acute toxicity is often used to give an indication of the usefulness of a food additive. Clearly, a substance exhibiting low acute toxicity would be a more likely candidate for a suitable food additive than would a substance showing appreciable acute toxicity. Subchronic and chronic toxicity studies should provide data on the cumulative effects, if any, of the test material. Feeding studies of 18 months in mice and 24 months in rats are considered adequate to ensure a minimum safeguard in evaluating the carcinogenic potential of a chemical additive.

WHO has been actively involved in promoting safe use of pesticides in both public health and agriculture (World Health Organization, 1975). WHO has conducted an intensive search for pesticides that could be used to control disease vectors and has conducted a "Programme for Evaluating and Testing New Insecticides" designed to locate materials that are not only effective against the target species but which are readily degradable and thus safe to use. A knowledge of the action of any new type of insecticide and its metabolism in mammals is used to assess the potential hazard to man. The Scientific Group on the Chemical and Biochemical Methodology for Assessment of Hazards of

Pesticides for Man recommended that WHO promote research on the metabolism and degradation of pesticides with special reference to new insecticides in order to elucidate the relationships between absorption, metabolism, and excretion of the parent compound and its degradation products in mammals.

Pesticides being considered for use in public health undergo thorough toxicological evaluations. Assessment begins during the early stages of the WHO Evaluation Programme with many otherwise suitable compounds being rejected almost immediately for toxicological reasons. Since skin contamination is the most important source of exposure, the dermal toxicity of a particular chemical substance often serves as the major factor limiting its further development. If a pesticide is to be used under conditions where drinking water might contain residues for a long period of time, special studies are carried out to assess its potential hazard to the general population.

Organization of Economic Cooperation and Development

The Organization for Economic Cooperation and Development (OECD) is comprised of an international group of 24 nations. The group was formed to ensure better health conditions as new products reach the international marketplace. Additionally, the founders hoped that some standardization of toxicological testing methodology would enable regulatory agencies of differing countries to work in concert (Organization for Economic Cooperation and Development, 1981).

The OECD has recommended four types of tests for inclusion in toxicity assessment:

1. acute, single species tests
2. chronic, single species tests
3. multispecies test systems
4. *in vitro* (suborganism) tests

Many of the specific testing procedures required by the EPA under TSCA are the same or similar to guidelines issued by the OECD. The reason for this is that the guidelines have evolved during the same time frame. Some of the methods required to carry out TSCA testing is obtained from OECD publications (e.g., Repeated Dose Toxicity Tests). A tabular summation of OECD testing guidelines follows (Table 10).

TABLE 10. OECD RECOMMENDED PROTOCOLS FOR TESTING NEW CHEMICALS

Type of Test	Species Used	#Animals per Dose Level	#Dose Levels	Duration of Test
Acute Oral Toxicity	Rat	10 (5F, 5M)	3	Single dose in a 24 hr. period, 14 day obs vation period
Acute Dermal Toxicity	Rat, Rabbit, Guinea Pig	10 (5F, 5M)	3	24 hr. exposure period, 14 day observation period
Acute Inhalation Toxicity	Rat	10 (5F, 5M)	3	4 hrs. exposure, 14 day observation period
Repeated Dose Oral Toxicity	Rat	10 (5F, 5M)	3 + control	14 or 28 days
Subchronic Oral Toxicity (Rodent)	Rat	20 (10F, 10M)	3 + control	90 days
Subchronic Oral Toxicity (Non-rodent)	Dog (Beagle)	8 (4F, 4M)	3 + control	90 days
Repeated Dose Dermal Toxicity	Rat, Rabbit, Guinea Pig	10 (5F, 5M)	3 + control (vehicle control if applicable)	21/28 days
Subchronic Dermal Toxicity	Rat, Rabbit, Guinea Pig	20 (10F, 10M)	3 + control (vehicle control if applicable)	90 days
Repeated Dose Inhalation	Rat	10 (5F, 5M)	3 + control (vehicle control if applicable)	6 hrs. exposure daily for 14 or 28 days
Subchronic Inhalation	Rat	20 (10F, 10M)	3 + control (vehicle control if applicable)	90 days
Chronic Toxicity	Rat, Dog, Primate	40 (20F, 20M)Rodents 8 (4F, 4M)Non-rodents	3 + control	12 months
Dermal Irritation/Corrosion	Albino Rabbit	3	1	4 hrs. exposure, 72 hrs. observation to maximum of 14 days
Eye Irritation/Corrosion	Albino Rabbit	3	1	24 hrs. exposure, 21 day maximum observation period
Skin Sensitization*	Guinea Pig	Dependent on test method chosen	Dependent on test method	Dependent on test method
Teratogenicity	Rat, Mouse, Hamster, Rabbit	20 pregnant rats, mice, hamsters, 12 pregnant rabbits	3 + control (vehicle control if applicable	Dose period covers the portion of major organogenesis
Carcinogenicity	Rat or Mouse	100 (50F, 50M)	3 + control	18-24 mos. for mice, 24-30 mos. for rats
Combined Chronic Toxicity/ Carcinogenicity	Rat or Mouse	100 (50F, 50 M)	3 + control	18-24 mos. for mice, 24-30 mos for rats

TABLE 10. (Continued)

Type of Test	Species Used	#Animals per Dose Level	#Dose Levels	Duration of Test
One-Generation Reproduction	Rat or Mouse	20 pregnant females	3 + control	
Two-Generation Reproduction	Rat or Mouse	20 pregnant females	3 + control	
Reverse Mutation Assay	*Salmonella typhimurium*	N/A	5 + 2 controls	48-72 hrs.
Reverse Mutation Assay	*Escherichia coli*	N/A	5 + 2 controls	48-72 hrs.
In Vitro Mammalian Cytogenetic Test	Mammalian Cells	N/A	At least 3 + 2 controls	Approximately 24 hrs.
Micronucleus Test	Mouse	10 (5F, 5M)	3 + 2 controls	72 hrs.

* Any one of the following test may be used: Draize, Freund's Complete Adjuvant, Mauer Optimization, Buehler, Open Cutaneous Test, Maximization, Split Adjuvant

From: Organization for Economic Cooperation and Development, 1981; Ibid, 1983.

STATE-OF-THE-ART OF DESCRIPTIVE AND MECHANISTIC TOXICOLOGY

For purposes of this discussion, descriptive, and mechanistic toxicology are differentiated (Klaassen and Doull, 1980). Descriptive toxicology deals directly with toxicity testing. It is primarily applied in an industrial arena to estimate the risk of a new product to man and the environment. It is essentially the segment of toxicology that is used to compare relative potency or hazard of the chemicals of interest in response to regulatory requirements. Mechanistic toxicology studies are designed to elucidate mechanism(s) by which a chemical or drug induces a toxic response and are generally research-oriented. Mechanistic toxicology investigations generate knowledge of basic physiologic and biochemical responses and can lead to advances such as the development of sensitive predictive tests for risk assessment. Descriptive toxicological studies, which incorporate standardized and commonly used test protocols, are conducted on a broad-scale basis by most laboratories. In contrast, mechanistic approaches often will reflect the particular industry (e.g., pharmaceutical needs and interests) and may not relate to testing requirement by a regulatory agency.

Descriptive Toxicology

A continued increase in the number and variety of chemical substances used in industrial, agricultural, and other applications coupled with the concern for public health has focused attention on toxicity testing. Although test requirements of different regulatory agencies are generally similar in sequence and composition, differences occur in the flexibility of approaches. The rigid approach to risk assessment is to require the same comprehensive battery of test procedures for all compounds, while the flexible approach is to require only those tests deemed necessary for a compound on a case-by-case basis. Flexibility in testing approach has a major influence on the cost/benefit structure of product development. The following material summarizes the principal components of descriptive toxicology, or toxicity testing, without regard to flexibility terms.

Toxicity testing includes acute, short-term (repeated-dose), subchronic and chronic toxicity assessments as well as special studies to investigate particular areas of importance (e.g., mutagenic, teratogenic, or carcinogenic potential).

Acute Toxicity

Acute toxicity tests, using single exposures, are of general importance as an initial screen for toxicity. One type of acute toxicity test, the LD_{50} (dose causing 50% mortality in a population) test, was developed in 1927 by J. W. Trevan for the biological standarization of dangerous drugs (Trevan, 1927). Acute toxicity tests are used to assess the relative toxicity of a substance. Additionally, the acute test can provide a first line of information on mechanisms of toxic effects.

Four tests are commonly conducted for acute toxicity: oral and dermal toxicity and skin and eye irritation (Chan et al., 1982). Inhalation toxicity and skin sensitization studies are often added, depending upon the nature of the test compound. The tests of choice are dependent upon the expected route of human exposure. An LD_{50} is determined in acute oral and dermal toxicity studies, whereas an LC_{50} (atmospheric concentration estimated to kill 50% of the animals exposed per unit time) is the unit of measure in acute inhalation toxicity tests. Skin and eye irritation studies are short-term tests which assess surface toxicities of liquids, solids, and aerosols to dermal and ocular tissues of animals. The final acute toxicity test mentioned, skin sensitization, is designed to observe immunologically mediated skin reactions to a test chemical. The observation period is generally one to two weeks, since skin sensitization is a delayed response. The improvement of methodologies used in assessing acute irritation and sensitization has been the subject of a considerable research thrust in recent years (Muller and Kley, 1982; Shultz and Fuchs, 1982).

The LD_{50} is not a biological constant, as the exact value will vary among species (Tattersall, 1982). Within a species the LD_{50} will differ with variables such as weight, sex, genetic composition, age, and diet. Therefore, except for under highly controlled conditions, the acceptance of LD_{50} values can be regarded as being relatively low. Despite these considerations, the LD_{50} is often determined with great accuracy, since a large number of animals can easily be tested in a short time to comply with various regulations and to assess relative toxicity. Although the determination of a precise LD_{50} may not be a necessary step in a chemical's safety evaluation (U.S. Food and Drug Administration, 1983), industrial laboratories often use the precision to 1) compare or rank the acute toxicity of various new substances and 2) satisfy the most stringent legislative requirements anticipated when a compound has a potential market in several countries (Zbinden and Flory-Roversi, 1981).

Short-Term Toxicity (14-28 days)

Short-term toxicity studies using repeated doses, are sometimes conducted to help further define the toxicity of a test compound. These studies provide additional information for the planning of long-term investigations and, as such, are used by industry to determine the appropriate dose range for subchronic or chronic studies. Oral, dermal, and inhalation exposure routes are all used in these assessments. The two-to-four week duration allows the observation of several endpoints, including target organ determination.

Subchronic Toxicity

Studies of intermediate length are often designed to be complete toxicological evaluations. Subchronic tests are conducted at lower exposure doses than acute tests and are designed to last approximately 10% of the test animal's lifespan (e.g., 90 d-rat; 360 d-dog). They provide appropriate information from which an evaluation of toxicity can be made and also provide information to facilitate the design of chronic studies. The subchronic study includes the observation of all effects under repeated administration of the test compound. This test regimen is suitable for the initial establishment of a no-effect level (NOEL).

The route of administration is the expected route of exposure for humans. Endpoints measured in the routine subchronic toxicity study include mortality, clinical chemistry, hematology, histopathology, and neurological and behavioral abnormalities. Reproductive toxicity data is now generated in some laboratories as a part of subchronic testing through the performance of such tests as sperm assays.

Results of subchronic tests are generally well accepted. These tests are considered essential for all new chemicals before any specific hazard level can be assessed.

Chronic Toxicity

Chemicals which are to be used in a manner that heightens the potential for chronic human exposure must be tested in a long-term exposure. The primary objectives of the chronic test are to identify and characterize the toxicity and to establish safe levels (NOEL) of the test substance. A well planned chronic toxicity assay includes the endpoints of mortality, food consumption, weight gain, hematology, clinical chemistry, urinalysis, organ weights, and histopathology (including tumorigenicity). The validity of the long-term assay is dependent upon the generation of good dose-response data for use in cross-species extrapolation.

A review of earlier material in this chapter serves to demonstrate that data submission requirements for compounds with different uses may vary considerably. The clinical signs of toxicity resulting from chronic exposure may only require 6 to 12 months. Some additional effects, such as the development of a cataract in the eye, may take a substantial period of time to become clinically detectable. Thus, the duration and other key parameters are established on the basis of previous short-term and subchronic test results. When an assessment of carcinogenicity is added to a chronic toxicity test, the study period will be lengthened to two-three years. However, due to worker and public safety concerns, the carcinogenesis bioassay has achieved a level of routine use in industrial toxicology laboratories. Table 11 lists the significant information surrounding in the two types of carcinogenesis evaluations now employed by testing laboratories.

TABLE 11. CARCINOGENESIS BIOASSAYS

Assay Type	Duration	Species	Information Generated	Endpoint
Carcinogenesis bioassay	24-30 months	2 required (rat, mouse)	Histology Gross pathology Immunology Hematology Preneoplastic changes	Carcinogenesis (based upon prior history of testing)
Transplacental multigeneration bioassay	36 months	2 required (rat, mouse)	Same as above + Reproduction	Carcinogenesis/reproduction (based upon prior history of testing)

Special Studies

When a compound elicits toxicities in acute and short-term tests which are indicative of particular areas of concern, such as developmental toxicity, follow-up studies are often implemented. Special studies are also required for particular types of chemicals (e.g., pesticides) and for chemicals which may have an impact on certain segments of the population (e.g., the aged).

These special studies may be performed as a part of subchronic or chronic toxicity tests or may be initiated as completely separate tests. Special studies primarily involve assessment of carcinogenicity, mutagenicity, reproductive toxicity (Table 12) or teratogenicity (Table 13).

Mechanistic Toxicology

Mechanistic toxicology includes the evaluation of hematology, clinical chemistry, urinalysis, organ and body weights, and/or gross and microscopic pathology endpoints as part of an acute, subchronic, or chronic toxicity study. It also encompasses special studies such as metabolic or pharmacodynamic studies, or evaluation of receptor binding, liver or kidney function, or reproductive endpoints. Those parameters which are most useful for the determination of mechanisms of toxicity depend to a great extent on the target organ for each toxicant studied. Generally, organ and body weights, hematologic indices, and organ function tests are useful parameters for assessing toxicity (Tables 14 and 15). Hematology, clinical chemistry, and liver and kidney function endpoints may contribute significantly to the estimation and elucidation of toxicity by indicating tissue or organ damage prior to the onset of observable structural changes (Falke, 1981; Heywood,

TABLE 12. REPRODUCTIVE TOXICITY TESTS

Test Type	Duration	Species	Endpoints			
One generation	6 months	1 required (usually rat)	Testicular evaluation Evaluation Sperm Fertility Male histopathology		Breeding success Litter size Pup survival Female histopathology	
Three generations*	18 months	1 required (usually rat)	Male histopathology Sperm evaluation Female histopathology Female evaluation	(F_0, F_1, F_2)	Body weight Litter size Reproductive efficiency	(F_0, F_1)
					Pup survival Development landmarks	(F_1, F_2)

*Used most frequently.

TABLE 13. TERATOGENESIS TEST

Duration	Species	Doses	Information Generated	Endpoints
4 months	2 required (rat, mouse, rabbit, hamster)	3 required (EPA: dose from implantation to C-section; FDA: dose from implantation through ogenesis	Fixed pups: skeletal malformations visceral abnormalities Fresh pups: Gross and visceral abnormalities	Number of corpora lutea Number of implantations Pul abnormalities Uterine weight Body weight (dam and pup)

Table 14. Standard Analytical Tests Associated with Hematology

Erythrocyte Count
Total Leucocyte Count
Differential Leucocyte Count
Hematocrit
Hemoglobin

TABLE 15. ORGAN-SPECIFIC TOXICITY TESTS

Type of Toxicity	Duration	Route of Administration	Enpoints
Pulmonary	91 days	Inhalation	Histopathology pulmonary function
Hepatic	>90 days	Oral, intraperitoneal	ALAT, ASAT, alkaline phosphatase, BSP retention, thymol turbidity, urinalysis, histopathology
Renal	>90 days	Oral, intraperitoneal	Blood urea nitrogen, creatinine, potassium, sodium, chloride, calcium, urinalysis, histopathology

Assays performed as part of subchronic and chronic toxicity tests.

1981). If the chemical acts on a particular target organ such as the central nervous system or reproductive tract, these general indicators of toxicity may not be sufficient to define any chemically-related deleterious effects and therefore may not be relevant for setting a NOEL.

Sex differences in response to chemical exposure are, in many cases, of limited importance when determining the approximate LD_{50} dose. However, differences in response between males and females or between pregnant and non-pregnant females becomes of greater importance in mechanistic evaluations. Chemicals that influence the endocrine system or reproductive tract, chemicals whose metabolic rate is altered by the hormonal status of the animal, and toxicants whose target organs are physiologically or metabolically altered by hormonal status or pregnancy are the most likely to demonstrate sex differences in toxicity.

Metabolism studies are of particular importance if the observed toxicity does not result from exposure to the parent compound, but rather from the toxicity of metabolites of the parent compound (Table 16). Extrapolation of data and calculation of risk to man is frequently complicated by different pharmacodynamics in different species. Many experiments may be needlessly done in inappropriate species whose metabolic handling of a chemical differs radically from that in man. Metabolism and distribution of xenobiotics should be considered early in the experimental hierarchy in order to better determine the appropriate species, route and rate of administration, and study endpoints.

Table 16. Metabolism Study

Duration	Species	Routes of Administration	Endpoints
8 Days	2-3 Required (rat, rabbit guinea pig, dog) Depending upon Compound	1) Always Intravenous 2) Expected Human Route of Exposure 3) Others (e.g., dermal)	Tissue Distribution and Retention Excretion Absorption Metabolite Characterization

Current Needs and Future Trends

Needs

The most obvious need remaining to be addressed in a comprehensive manner is a critical review of current testing procedures. Once such a task is completed, the concepts favoring reduction, expansion, or the status quo can be tested. The National Academy of Sciences has recently published the results of a study conducted for the NTP on strategies to determine needs and priorities in toxicity testing (National Research Council, 1984). This material is a good foundation for further work on the subject. The study group divided a select universe of compounds into 7 categories in order to characterize toxicity testing:

1. Pesticides
2. Cosmetic ingredients
3. Drugs and excipients
4. Food additives
5. Chemicals in commerce (at least 1 million pounds per year)
6. Chemicals in commerce (less than 1 million pounds per year)
7. Chemicals in commerce (production unknown)

It was concluded that enough toxicity and exposure data to make a complete health hazard assessment are available for only a small fraction of the several hundred compounds assessed. A significant amount of useful information generated by descriptive toxicologic approaches was lacking. A general conclusion is that tests of some chemicals will always be necessary, but the most time- and cost-efficient means of approaching future toxicity testing will be to gain understanding of mechanisms at whole animal, organ, cellular, and subcellular levels. This will result in an increased emphasis in the area of mechanistic toxicology. These considerations, combined with retrospective studies on historical compounds, must be made prior to any efforts to change the status of toxicity testing.

Whether or not change is deemed a necessity, the following points should be addressed. Efforts should be made to incorporate emerging techniques, such as immunochemistry, into routine toxicological investi-gations. Valid screens need to be established for the important areas of neurotoxicity and various aspects of reproductive toxicity. In addition, emphasis must be placed on cross-species extrapolation. Ramifications of the introduction of a requirement for a second test species under given circumstances should be considered. The above points all aim at improvement of the integrity of current hazard assessment programs.

Trends

The visible trends in toxicology include changes in practices and protocols and the incorporation of existing concepts and alternative technologies. The Good Laboratory Practice (GLP) procedure requires that all laboratories conducting toxicological studies ensure that these studies are carried out according to an acceptable high standard and involve a comprehensive auditing system. The advantages of such a system in terms of data collection, processing, and reporting are rather obvious. One drawback of GLP is associated rise in time and costs involved with these studies. A future trend will likely be the reduction of GLP procedures during phases of toxicity testing which do not affect regulatory decisions (e.g., short-term toxicity studies).

A tiered approached to toxicity testing is gaining acceptance (Weisburger and Williams, 1981). A tiered, decision point approach is the most logical means of assessing toxicity, while pinpointing the necessary special studies, subchronic studies, and chronic studies to run. Modifications of existing test protocols (e.g., reducing the number of animals killed in an acute toxicity study) will be fashioned into the tiered, decision point approach.

Incorporation of existing concepts into testing strategies will undoubtedly occur. The increased use of metabolism and pharmacokinetic studies prior to initiation of subchronic or chronic toxicity evaluations will aid the decision-making process. The area of mechanistic toxicology should expand as more emphasis is placed on mechanisms of action, molecular-biological interactions, and quantitative structure-activity relationships. An additional area, which has already gained considerable attention, is the assessment of behavior in toxicity studies. Numerous neurobehavioral and behavioral toxicology techniques (e.g., running wheel and residential maze) are available and some are presently in use. This area is rapidly developing but has not yet gained widespread acceptance among toxicologists.

The final emerging trend is the incorporation of alternative technologies to the use of traditional mammalian models. Mutagenicity testing schemes are already actively validating and employing non-vertebrate organisms. The use of alternative technologies will continue to increase. This is the subject of the next section of this chapter.

ALTERNATIVES TO ANIMAL USE IN TOXICITY TESTING

ALTERNATIVE TECHNOLOGIES

Novel approaches to toxicity testing are currently the subject of a considerable number of investigations. These approaches include simple modifications of existing test formats, the incorporation of adjunct tests, and the use of new test systems. In order to summarize activities in this area for the reader, this section is divided into mammalian and nonmammalian approaches. Mammalian approaches are further subdivided into (a) modified whole animal, (b) *in vitro* culture, and (c) noninvasive human methodologies. The nonmammalian approaches consist of (a) vertebrates, (b) invertebrates, (c) microbes, (d) plants, and (e) computer models. Each of these approaches is generally considered in the following text.

Mammalian Approaches

Modified Whole Animal Approaches

Many endpoints in toxicological research cannot be addressed adequately with *in vitro* systems, but rather require whole animal *in vivo* experimentation. These endpoints include, but are not limited to, the evaluation of behavior, sensory function, delayed effects, and anticipated physiologic toxic responses in animals and humans. Several modified whole animal approaches have been recommended to reduce the number of animals used or reduce the animals' discomfort.

One suggested approach is a sequential study design. In a sequential study, the same animals receive increasing doses over a period of time in contrast to the more conventional parallel design with a fixed number of animals at each different dose level run simultaneously. The advantage to a sequential study design is that the necessary number of animals is greatly reduced. This method can be used to approximate the LD_{50} and provide acute toxicity data. A disadvantage is that it does not take into account either cumulative dose-effects or delayed toxicity.

The use of historic control data has been suggested as an alternative method for reducing the number of animals used by eliminating the necessity for concurrent controls. Routine test protocols include a control dose group run simultaneously with experimentals to ensure that treatment effects are discerned. The difficulty with reducing the number of controls is that the conditions of a different study can never be exactly duplicated for sex, age, strain, weight, and health of the animals as well as season, environmental conditions, and animal treatment. Randomization of animals between treatment and control groups would no longer be possible, so assurance of similarity between groups would be lost. However, if the historic data base is sufficiently large, the significance of interanimal and interstudy variations would be maximized. Under these conditions historic data would represent a reasonable base for comparison to test groups and might be used to eliminate the need for concurrent controls.

Other recommendations for modified whole animal toxicity studies include reducing the number of animals in each dose group and increasing the amount of toxicity data gathered from each animal. Thus, in lieu of an LD_{50} study or an acute toxicity study with few study endpoints, an acute toxicity study with clinical chemistry, hematology, physiologic (clinical) observation, behavioral, pathologic, and histopathologic endpoints would be conducted.

Likewise, chronic one year toxicity studies with many toxicologic endpoints, serial sacrifice providing information on the etiology of disease, and attention to preneoplastic changes may be of greater utility than chronic two-year bioassay carcinogenesis studies which may be conducted with twice the number of animals.

A proposed method to reduce the discomfort of the animals is the use of analgesics during testing. This may be impractical in instances where increased or decreased analgesia is an anticipated behavioral response to the toxicant or when analgesic treatment potentiates or diminishes the toxic effects of the test compound. However, when it is known that analgesic treatment does not interfere with test responses the use of such tactics should be encouraged.

In Vitro Culture Techniques

In recent years, the primary focus of alternative toxicity research has been on *in vitro* culture techniques. *In vitro* systems range from bacterial cultures (see Microbes) to cell-free systems, to cell, tissue, organ culture or tissue slice culture, to rodent whole embryo culture. Advantages of *in vitro* systems include:

- reduced numbers of animals required
- the possibility of conducting numerous simultaneous replicates
- the use of human tissues or cell lines, and
- the short duration of the test times.

Disadvantages include questionable extrapolation between phylogenetically distant systems (i.e., bacteria to man), a non-physiologic milieu, and the inability to reproduce complete physiologic responses (i.e., as when in vivo the response to the toxicant includes (a) metabolic activation/deactivation at a distant site; (b) the recruitment of macrophages or lymphocytes; or (c) a complex inter-action of several tissue types, e.g., when primary eye irritation occurs).

The use of in vitro culture techniques can be divided into two basic genres: those used for assessing mutagenicity/carcinogenicity and those used for evaluation of other toxicities. DNA damage is viewed by many as a common link between mutagenic and carcinogenic events. However, DNA damage is not the only currently accepted theory for cancer induction. It is now generally accepted that no single short-term test can be relied upon to detect all classes of mutagens and carcinogens. A battery of genetically and phylogenetically distinct short-term tests should be conducted on a new chemical prior to the classification of its mutagenic/carcinogenic potential. A tiered system of assays is most commonly employed. Commonly used in vitro mutagenicity/carcinogenicity test systems are listed in Table 17.

Cytotoxicity. In vitro cytotoxicity tests may use cell viability or cell reproduction as indicators of general toxicity or may be used to describe specific toxicities such as neurotoxicity or hepatotoxicity. Cell culture systems used for neurotoxicity include dissociated embryonic mouse spinal cord, neuroblastoma, dorsal root ganglia, and co-cultures of spinal cord and muscle. Tissue culture techniques are currently being developed for the in vitro testing of chemicals for irritants using skin cultures and for ocular toxicity to replace the Draize rabbit eye test. A variety of ocular cell types have been cultured: corneal epithelium, fibroblasts, endothelium,

TABLE 17. *IN VITRO* MAMMALIAN TEST SYSTEMS

1. Cell transformation

    1.1. Cultured baby Syrian hamster kidney cells (BHK 21)
    1.2. Cultured mouse fibroblasts (BALB/C 3T3)
    1.3. Cultured mouse embryofibroblasts (C3H 10T 1/2)
    1.4. Cultured Syrian hamster embryo cells
    1.5. Viral enhancement in cultured Syrian hamster embryo cells (SHE SA7)
    1.6. Viral enhancement in cultured rat embryo cells

2. Mammalian cell cytogenetics in vitro

    2.1. Chromosomal aberrations
    2.2. Sister-chromatid exchange

3. Point mutation system in cultured mammalian cells

    3.1. Hypoxanthine-guanine phosphoribosyl transferase (HGPRT)
    3.2. Thymidine kinase (TK)
    3.3. Oubain resistance

4. Mammalian cells in culture - epigenetic systems other than cell transformation

    4.1. Chinese hamster fibroblast (V7G) primary cell co-cultivation
    4.2. Inhibition of metabolic cooperation

5. Assays based on estimations of DNA interactions

    5.1. Unscheduled DNA synthesis *in vitro*
    5.2. DNA damage - alkaline sucrose gradient sedimentation analysis

retinal pigment epithelium, and lens epithelium. Representative endpoints for *in vitro* cytotoxicity systems are listed in Table 18.

Organ/Embryo Cultures. Additionally, whole embryo cultures or organ cultures can be used and may represent the closest *in vitro* approximation to the *in vivo* whole animal model. Embryo cultures can be used as an alternative mammalian system for the evaluation of teratogenic potential and may also have limited use in indicating potential target organs for later toxicity studies. Organ cultures of perinatal tissues can be used for morphogenesis and endocrinology studies, but most mammalian adult tissues show limited survival in culture. Tissue slices and whole organ perfusions are techniques most commonly used with adult organ systems, but of necessity, these methods are of very short duration. Organ cultures may be useful for metabolism or pharmacodynamic studies and can be used to examine drug interactions, mechanisms of chemical effects, and chemical biotransformation.

Non-Invasive Human Methods

Chemical injury can result from a variety of routes of exposure including absorption through the skin. Substances that come in contact with the skin and eyes can also produce local effects at the site of contact. Draize et al. (1944) initially developed and applied quantitative methods to skin and eye tests by using an arbitrary weighted scale for describing the degree of the various effects that are observed. The modern tests (Vinson et al, 1969, Harber and Baer, 1969; Conquest et al., 1977) still involve the principles established by these workers.

TABLE 18. REPRESENTATIVE ENDPOINTS FOR IN VITRO CYTOTOXICITY TEST SYSTEMS

1. Cell viability
    - (a) Vital staining-dye exclusion
    - (b) Release of radioactive substances
2. Cell reproduction
    - (a) Growth of high-density cultures
    - (b) Cloning efficiency
3. Cytological anomalies
    - (a) Light microscope level
    - (b) Ultrastructural level
4. Differentiated function
    - (a) Collagen synthesis
    - (b) Hormone production
    - (c) Cell differentiation
    - (d) Immune response
    - (e) Cell surface-receptors
5. Balanced biosynthesis
    - (a) DNA levels
    - (b) RNA levels
    - (c) Protein levels

Dermal. In the human, local dermal effects that can occur are: (1) primary irritation, (2) corrosion, (3) cutaneous sensitization, (4) phototoxicity, and (5) photoallergy. Limited screening tests on animals are necessary before tests on human skin are contemplated, although the need for these should be obviated with the development of reliable *in vitro* techniques. The animals that have been widely used for the detection of irritant properties of chemicals are the albino rabbit, the albino guinea pig, and the albino mouse. Evaluation of skin effects involves using a scoring system which has been published in the Federal Register in Section 191.11 of the Federal Hazardous Substances Act of the United States. The scoring involves relative assignment of separate numbers for the degree of erythema and the degree of edema. Six to twelve animals are used for the acute dermal 3-day test. Factors that may influence the results of skin tests on man are:

1. Individual variation in susceptibility to cutaneous irritation
2. Ambient condition
3. Season of year
4. Anatomical site chosen for test
5. Age
6. Concentration of test substance
7. Vehicle employed
8. Method and frequency of application to skin

Keeping these factors in mind, conditions for a 'standard test' methodology can be established. *In vitro* techniques are potentially of great value for many aspects of skin toxicity testing, but, with the exception of those that relate to phototoxicity, are still at the developmental stage. The evaluation of stratum corneum barrier function using cell culture methods presents obvious difficulties because in vivo interacting factors such as nerves and vessels are absent. For this purpose whole skin taken either from cadaver or from operation specimens can be used. Useful methods that are being developed include: (1) a chamber technique for the estimation of respiratory enzyme release from the skin and (2) measurement of enzyme activities in suction blisters.

Ocular. A reliable assessment of potential chemical hazard to the eye is necessary because of the great psychological and physical significance of visual damage, impairment, or loss in man. To date no in vitro methods have been developed that would replace the in vivo model for final safety clearance or hazard evaluation. Draize (1959) developed a subjective procedure for evaluating the irritant capacity of liquids and solids. At least six rabbits are used in a single study. The degree of local reaction is assigned a numerical score according to an illustrated color photographic guide of rabbit eyes for grading eye irritation. Recently, a number of workers in this field have developed an objective methodology for evaluation of ocular toxicity (Conquest et al, 1977). The tests involve measurement of corneal thickness, the corneal and conjunctival water content, and permeability of the blood vessels in the eye. Research in this area should emphasize in vitro techniques: (1) chemical tests for binding or denaturation of enzyme systems and/or inactivation of enzyme systems concerned in tissue homeostasis (i.e., develop systems to detect the loss of regulatory ability), (2) validation of techniques employing isolated eyes, (3) primary cell and organ cultures of eye-derived tissues and (4) general cell culture lines and strains.

Nonmammalian Approaches

Vertebrates

While several validated tests exist which use nonmammlian vertebrate organisms (primarily fish and birds) in the assessment of environmental toxicity, their potential role in the assessment of human health hazards has yet to be ascertained. Due to increased interest in finding suitable alternatives to traditional mammalian assays, a significant number of test systems are currently being developed which use a wide array of fish, amphibian, and reptilian species.

In the area of carcinogenesis, alternative tests utilizing various fish species (e.g., rainbow trout, medaka, sheepshead minnow, bullhead) represent three approaches to determining potential carcinogens: (1) the long-term bioassay, (2) the short-term screening test, and (3) the use of indicator species in field monitoring studies. These tests include fish whole animal bioassays, the mixed-function oxygenase activity assay, and the routine sampling of natural fish populations in the laboratory and field for tumorigenesis.

Fish are also used in certain mutagenesis assays. The fish chromosome damage assay involves the application of the sister chromatid exchange assay (SCE) to the study of chemical mutagens in fish. Fish cell cultures may also be used to detect chemical mutagens. A large number of cell lines are commercially available and screening tests for toxicity and mutagenesis can be performed in a short time frame.

Nonmammalian vertebrate test systems have been developed in the area of teratogenesis because no validated screening tests currently exist. An effort to validate teratogen prescreens has not been undertaken by the NTP. These systems are based upon the concept that early developmental stages are mechanistically comparable among the various animal phyla. The chick embryo has received considerable attention for this purpose. One of the most promising assays developed is the FETAX: Frog Embryo Teratogenesis Assay: *Xenopus* which uses the South African clawed frog, *Xenopus laevis*. This assay is rapid (96 hrs.) with well-defined endpoints and yields excellent dose-response curves. Similar analyses can be performed on reptilian embryos, however, amphibians are the preferred test species because they are easier to obtain and culture and have been subjected to more intensive investigation than reptiles. Chemical teratogens have been shown to cause severe morphological abnormalities in fish. The establishment of well-defined endpoints has resulted from a significant amount of research on various fish species; however, there are significant differences in sensitivity between species. A good model fish species for use in screening teratogens remains to be found.

Reproductive dysfunction in nonmammals could be predictive of a similar hazard in mammalian systems. Measures of impotence, sterility, fecundity, fetal loss, and infant mortality can be completed with many nonmammalian species. Little comparative information is available on this concept. Reproductive and embryotoxicity testing with amphibians (primarily *Xenopus*) and with fish could be important subchronic and chronic alternative tests for the purpose of screening chemicals and environmental monitoring.

Few adequate tests currently exist in the field of neurotoxicity. Various screening tests for sensory and motor inhibition, receptor level inhibition, altered behavior, and behavior could be developed with further research on alternative nonmammalian species. The use of birds in the study of esterase inhibition has expanded our knowledge of neurotoxic effects of pesticides and related substances.

When a chemical affects toxicity in a certain organ or organs, this would suggest performance of an organ-specific toxicity study. In these cases, nonmammalian vertebrates could be used in screening tests for subchronic toxicity. Potential screening tests for hepatotoxicity (liver) using various fish species (e.g., rainbow trout) and renal (kidney) toxicity using fish (flounder, mullet, rainbow trout, etc.) and amphibians (frog, toad) show promise.

The fate of a chemical substance after it enters the body ultimately determines its toxicity. An understanding of a chemical's metabolic fate is extremely valuable in assessing its impact on environmental health. Studies of metabolite formation in amphibian and fish species could conceivably forecast potential problems in man.

Invertebrates

Although widely used in assessment of environmental toxicity, invertebrate organisms have not been as extensively used in toxicity testing for human health effects. However, due to the recent search for alternatives to mammalian testing, a large number of assays have been developed which utilize invertebrate species from several different phyla.

Perhaps the best known assays involving invertebrates are the various mutagenesis assays which use the fruit fly, *Drosophila melanogaster*. Several tests with fruit flies have proven valuable in understanding chemical mutagenesis. These include the Sex Linked Recessive Lethal (SLRL) Test, Heritable Translocation Test, Chromosome Loss Test, Nondisjunction Test, and

Specific Locus Test. Among these, the SLRL Test and the Heritable Translocation Test have been validated and have broad application. The SLRL detects lethal effects for a number of types of genetic damage, whereas the Heritable Translocation Test is specific in detecting chromosome breakage and rearrangement. *Drosophila* genetics are well-understood and the organism has already gained widespread use as an alternative species. In addition to its use in mutagenesis studies, *Drosophila* also has potential for use in teratology and reproductive toxicity studies.

Various protozoans have been used in toxicity testing, primarily in the assay of cigarette smoke. They are currently being proposed for use as alternative tests for carcinogenesis, mutagenesis, and reproductive toxicity.

The sea urchin has long been a favored test organism for basic reproductive research. Consequently, the mechanisms and procedures are familiar to many biologists. The sea urchin model for fertilization and development can be used in screening for reproductive toxicity, teratogenicity, and mutagenesis.

Currently, no validated screening tests are available for detecting potential teratogens. Consequently, much focus has been directed towards this area of toxicology. This has resulted in the development of a number of vertebrate and invertebrate test systems. Among invertebrate organisms used for these teratogenic assays are brine shrimp, hydra, planaria, insects, and sea urchins. Hydra has received considerable attention as a tool for studying basic mechanisms of development. The specificity and sensitivity of this test system has fared well in studies with known mammalian teratogens.

A definite need exists for the development of adequate tests for neurotoxicity. Aside from testing for overt neurotoxicity, few assays exist which employ invertebrates.

Nematodes, annelids, and molluscs are also employed in alternative mutagenesis testing regimes and, additionally, molluscs are used in the area of reproductive toxicology. Invertebrates (sponges, molluscs, crustaceans, echinoderms) are also being used as test subjects in metabolism studies. Studies of metabolite formation in nonmammalian species can lend insight to interspecies variation observed in mammals and generate information on mechanisms of toxicity.

Microbes

In recent years much emphasis has been placed on the utilization of bacteria and fungi as indicators of toxic drug effects in higher organisms, including man. There are many advantages of using such single cell organisms for toxicological testing. Organisms can be cultivated in large numbers and screened concurrently. Results can be obtained in a much shorter period of time than in traditional mammalian test systems and the cost of conducting assays is significantly reduced. A number of validated microbial test systems exist and are considered below.

A number of bacterial and fungal species are currently used as quick and effective screens for specific types of toxicity, e.g., the Ames *Salmonella* test, *E. coli* reverse mutation assay, *Aspergillus* and *Neurospora* gene mutation assays, and the *Saccharomyces* gene conversion test in mutagenesis screening test batteries (Malling, 1978). The use of bacteria as indicators of chemically induced mutagenesis has gained widespread importance with the development and validation of the Ames Salmonella test. Bacteria offer certain advantages over more complex entities, namely:

- a large number can be treated and screened in a short period of time,
- they are haploid organisms with small genomes and short generation times,
- they can be grown on a defined media,
- they can be used in the detection of forward and reverse mutations,
- they can be used in the detection of both mutagenic and inactivating DNA alterations.

Fungal systems have been shown to be highly useful in mutagenicity testing because they indicate the relative potency of mutagens. In many instances, they are more sensitive than prokaryotic systems--a number of mutagens have been detected which tested negatively in bacterial systems (Zimmermann, 1983). Despite this, the fungi do not offer more variety in testing then bacteria. Their importance to mutagenesis testing is in the first tier of a sequential test format.

Plants

Utilization of plants to predict toxicity in mammalian systems is neither well-developed nor well-established. When attempting to use plant tissue for toxicity testing of chemicals and when extrapolating these results to mammalian systems, the inherent biochemical and structural differences between the two must be considered (Woodard, 1976).

Testing with plant species has been used primarily in the area of mutagenesis. Experimental work has been done with the spiderwort, *Tradescantia* (detection of somatic mutations), and with corn, barley, onion, and bean (detection of chromosomal aberrations). A promising mutagenic assay uses maize (*Zea mays*) in a test designed to take advantage of unique genetic markers of the maize plant by comparing plant mutagenesis with plant activation of standard bacterial mutagenesis assays. Previously undetectable mutagens have tested positively with this system.

Computer Models

Advances in computer technology during the past twenty years have played a major role in the increased development of sophisticated mathematical models of structure-activity relationships. These models are used to predict biological response on the basis of either a physical or chemical property or structural index. Models frequently take the form of an equation in which the biological response is expressed logarithmically and correlated with indices of molecular structure or property. The indices generally represent degrees of variation of a specific property among a series of related chemicals. If such an equation, known as a "QSAR" (quantitative structure-activity relationship), is found to be statistically significant, it then becomes a useful tool.

A significant correlation with physical, chemical, or structural indices can be used to predict the biological response of a series of related compounds and can also provide valuable insight into the mode of action of a set of closely allied compounds. Since a QSAR has both predictive and diagnostic utility, there are two primary ways in which it could be used in limiting animal experiments:

1. If QSAR results indicate that a series of compounds should no longer be synthesized in a particular manner, then conceivably no more compounds of this nature would be produced and no further resources, including animals used in toxicity studies, would be sacrificed.
2. If the QSAR predicts that an available but untested compound will cause a certain response in an animal and provided there are not regulatory agency requirements for animal experimentation, then animals will not be used. The response predicted by the QSAR may be beneficial, it may indicate an overall toxicity (LD_{50}), or it may indicate a specific toxic effect, such as mutagenicity (Tute, 1982).

According to the International Union for Pure and Applied Chemistry working party on QSAR, the following three criteria should be met in order to develop a QSAR capable of predicting a biological end-point.

A. The mechanism by which the response occurs should involve a common rate-determining step

B. The compounds considered should belong to a closely related series

C. The active site in the biological system should be the same for each compound.

However, Enslein et al. (1983) discuss evidence that conditions B and C are not necessary. Further research in this area must be completed in order to establish the steps necessary to develop a QSAR, as even analysis of diverse structures may prove to be of great value (Bawden, 1983). Because lethal toxicity can result from the action of many mechanisms and can involve several target organs and enzyme systems, the derivation of a model capable of predicting acute toxicity as expressed by LD_{50} presents a challenge to the modeler. The difficulty in deriving such a model lies in conditions A and B not being met. In addition, the site and mechanism of action is often unknown or ill-understood. In spite of this, models for predicting acute toxicity have been proposed but presently cannot replace animal experimentation. QSARs for predicting mutagenicity and carcinogenicity have also been developed. Mutagenicity and carcinogenicity data appears to be more easily describable mathematically than does LD_{50} data. Examples of each mathematical model type are cited in Table 19.

Many research laboratories, especially in the pharmaceutical industry, have incorporated mathematical modeling into their research programs. Since pharmacotoxicity shares a close relationship with pharmacokinetics, the pharmaceutical industry has used mathematical models to obtain as much information as possible from its pharmacokinetic studies.

The ultimate goal of mathematical modeling is to be able to assess the potential beneficial or harmful effects of an untested chemical substance from a basic understanding of its structure.

INCORPORATION OF ALTERNATIVE TECHNOLOGIES INTO EXISTING TESTING REGIMENS

Currently used toxicity testing schemes have taken years to develop and validate. It should be apparent that animal testing cannot be eliminated in the foreseeable future (Fund for the Replacement of Animals in Medical Experiments, 1983). However, mammalian testing strategies are varied, have somewhat less than precise components, and typically require substantial amounts of time to conduct at significant costs.

TABLE 19. EXAMPLES OF MATHEMATICAL MODELS

Acute Toxicity	Rat LD_{50}	2-trifluoromethylbenzimidazoles	Adamson, Bawden, Saggers (from Tute, 1983)
Acute Toxicity	Rat LD_{50}	1851 compounds from the Toxic Substance List	Enslein et al. (1983)
Acute Toxicity	Acute Toxicity of Hypnotics and Sedatives	Barbiturates	King & Moffatt (1981)
Acute Toxicity	Mouse LD_{50}/Worm Behavior	Alcohols	Tichy & Trcka (1983)
Mutagenicity	Electron affinity/ mutagenic potency	Nitro imidazoles Nitro benzenes Nitro furans	Chin et al. (1978)
Mutagenicity	Mutations in *Salmonella typhimurium*	Triazenes; platinum complexes	Hansch et al. (1980)
Mutagenicity	Mutations in *S. typhimurium*	472 compounds	Enslein et al. (1983)
Carcinogenicity	Pattern recognition/ discriminant analysis/ chemical carcinogens	N-nitrosoamides and amines	Chou & Jurs (1979)
Carcinogenicity	Pattern recognition/ discriminant analysis/ chemical carcinogens	Nitrosamines	Rose & Jurs (1982)
Carcinogenicity	Pattern recognition/ discriminant analysis/ chemical carcinogens	343 compounds	Enslein & Craig (1982)
Teratogenesis	Statistical Structure-Activity Model/Teratogenicity	430 compounds scored as to their teratogenic potential	Enslein, et al. (1983)

The criteria for determining alternative technologies which have significant potential to refine, reduce, or replace currently used toxicity tests center initially on those same criteria which are the basis for successful mammalian models, i.e., sensitivity, specificity, replicability, reliability, and degree of extrapolation. In addition, an alternative test should demonstrate improved features such as reduced time to endpoint, decreased cost, and perhaps an improved performance as compared to existing tests. Thus, for the ideal alternative the response to toxicants should be similar to that of man and the substance in question should be handled in a manner similar to that of man. Factors to be considered are uptake, metabolism, retention, excretion, and pertinent physiological possesses (e.g., circulatory function). A comparative relationship should also exist between the sensitivity to toxic substances in the alternative and man. Generally, degree of sensitivity and similarity in metabolism are the chief factors influencing species choice for a safety evaluation, but this is not always the case (Oser, 1981). In some cases, a particular species is chosen simply because of a researcher's familiarity with it.

Although the tendency is to use an organism most closely related to man on a phylogenetic basis, no species will resemble man in all respects. A high degree of variability in toxicologic response can occur between closely related species. Physiologic similarities (as homologues and analogues) can be most useful in hazard asessment. For example, the lower vertebrates appear to physiologically handle many trace metals in a manner similar to that shown by the higher vertebrates. Homologous organ systems often exhibit nearly identical pathways of metal metabolism.

An alternative test system can still be used in safety evaluations, even if some of the above criteria are not adequately addressed. Although they may not generate the total information necessary to make definitive regulatory decisions, these models can serve to screen and prioritize chemical substances. A specific example of an alternative test system which has proven superior to and actually replaced an existing test is the *Limulus* amoebocyte lysate test which replaced the rabbit pyrogen test in screening for endotoxins.

Testing schemes for virtually all major toxicologic endpoints have potential for refinement, reduction, and/or replacement to enhance the models' sensitivity and specificity. In the case of carcinogenicity, mammalian testing requires large number of animals and is characteristically long (due to the animals life span) and costly. For mutagenicity, various alternative mammalian and non-mammalian (e.g., Ames *Salmonella* and *Drosophila* SLRL) tests are already being utilized in genetic toxicology test batteries. Additional tests for mutagenic properties such as gene translocation are needed as well as tests for mutagenic potency. Issues surrounding reproductive toxicity include the cost and time required for current testing strategies and the need for short-term screening tests. For teratogenicity, there are currently no validated screening tests. There is a rapidly growing level of interest developing for both teratogenicity and neurotoxicity in assessing potential health impacts. There are no good tests currently in existence which adequately assess the neurotoxicity of compounds other than organophosphates (e.g., pesticides, certain industrial and military compounds). Due to the complexity of the nervous system, biochemical/pharmacological endpoints are presently of limited use. Consequently, the assessment of neurotoxic effects is focused on behavioral and functional tests. Both neurotoxicity and teratogenesis have significant potential for application of alternative species testing schemes.

Based upon the already demonstrated worth of alternative test systems which are currently being used (e.g., Ames *Salmonella* and *Drosophila* SLRL) and the potential new systems identified in this study, it should be apparent that alternative technologies can play an increasingly important role in the future of human health and safety evaluations.

Validated Alternative Test Systems

An alternative test system must undergo a rigorous period of validation in order to gain acceptance in scientific and regulatory arenas. It takes several years to develop a new assay to the point where it becomes a viable component of a testing regime. A number of lower life form or alternative test systems have been subjected to an extensive and thorough validation process and are currently being utilized in testing for specific toxicities (U.S. Environmental Protection Agency, 1978). Table 20 presents a summary of validated alternative systems presently used in genetic toxicology.

Alternative Toxicity Test Systems Under Development

A large number of test systems are currently under development. In retrospect, it takes approximately 20 years from the time of inception of a toxicity assay system until its validation and acceptance for use in human safety evaluations (Sabourin and Goss, 1984). Therefore, at this point in time we can assume that numerous systems are undergoing the evaluation process. Tables 21 and 22 list some representative mammalian and nonmammalian test systems, repsectively, which are currently being developed and/or validated. The toxicity endpoint, test endpoint, and most commonly used organism are supplied for the convenience of the reader. More information on these systems can be obtained using the references given. These lists are by no means complete. However, they do represent the majority of test systems which can be incorporated into toxicity testing programs in the near-term.

TABLE 20. REPRESENTATIVE NONMAMMALIAN ALTERNATIVE TOXICITY TEST SYSTEMS

Test Type	Test Endpoint	Test Organism	Reference
MUTAGENESIS			
Plant-microbe bioassay	Gene reversion	Plant/microbe	Gentile and Plewa (1983)
Tradescantia stamen hair	Somatic mutations	Plant	Underbrink et al. (1973)
Protozoan mutagenesis assay	Significantly reduced vigor and increased mortality in progeny	Invertebrate	Smith-Sonneborn (1983)
Mutagenesis in sea urchin embryos	Mitotic aberrations, micronucleus formation	Invertebrate	Hose an Puffer (1983); Pagano et al. (1982)
Mutation in fish cell cultures	Mutant colonies	Vertebrate	Kocan et al. (1979, 1981)
TERATOGENESIS			
Insect embryo culture assay (fruit fly)	Muscle/neuron differentiation	Invertebrate	Bournias-Vardiabasis et al. (1983)
Sea urchin embryo assay	Developmental abnormalities	Invertebrate	Hagstrom and Lonning (1983)
Hydra reaggregation	Dissolution of embryo	Invertebrate	Johnson and Gabel (1982)
Amphibian embryo assay	Developmental abnormalities, skin pigmentation, size differences, survival	Vertebrate	Browne and Dumont (1979)
Chick embryo assay	Malformations	Vertebrate	Jelinek (1982)
CARCINOGENESIS			
Photodynamic assay	Cell immobilization/death	Invertebrate	Smith-Sonneborn (1983)
Fish whole animal bioassay	Tumorigenesis in liver tissue	Vertebrate	Bailey et al. (1982); Koenig et al. (1982)
REPRODUCTIVE TOXICITY			
Echinoderm fertilization and development	Gamete viability; fertilization success; developmental abnormalities	Invertebrate	Greenwood (1983)
Insect reproduction	Failure to hatch, mortality, developmental abnormalities	Invertebrate	Srinivasan and Kesavan (1979)
Fish reproduction	Mortality of eggs and sperm, failure of eggs to hatch, oran abnormalities	Vertebrate	Bresch (1982)
NEUROTOXICITY			
Neurobehavior coupled to function	Behavioral, physiological	Vertebrate/ invertebrate	Haynes (1980); Hill et al. (1981)
Chemoreception	Magnitude, frequency, and latency of electrical response	Vertebrate/ invertebrate	Hara (1979)
Nerve regeneration	Nerve regeneration	Vertebrate/ invertebrate	Davis and Benloucif (1981)

TABLE 20. (Continued)

Test Type	Test Endpoint	Test Organism	Reference
NEUROTOXICITY (Continued)			
Receptor binding assays	Differences in specific binding of pharamocological agents after incubation of neuronal membranes in presence and absence of a competing chemical (% of ligand bound to receptor and amount bound)	Vertebrate	Bondy (1982)
Cholinergic screening assays	Acetylcholinesterase (AChE) activity; AChE-inhibitor tolerance; I_{50} for AChE (50% reduction of activity)	Vertebrate	Olsen and Christensen (1980)
Nervous system histopathology	Morphology of brain and eye	Vertebrate	Hawkes and Stehr (1982)
ORGAN-SPECIFIC TOXICITY			
Liver function	Bromosulphothalein clearance, serum alanine aminotransferase activity (ALAT), seram aspartate aminotransferase activity (ASAT), liver glycogen, liver lactate, serum glucose	Vertebrate	Weber et al. (1979)
Kidney function and transport	Ion transport, ATPase activity, serum creatinine, serum magnesium, serumcalcium, serum phosphate, pathologies	Vertebrate	Pritchard and Miller (1980)
METABOLISM			
Cell/organ culture methods	*In vitro* toxicity; specific responses physiological substances; mutation frequency; enzyme activity; protein, RNA, DNA content; glucose release	Vertebrate/ invertebrate	Balls et al. (1982); Marion and Denizeau (1983)
Amphibian metabolism	Metabolism of parent compound (rate and amount), metabolites formed and amount	Vertebrate	Clothier et al. (1981)
Fish metabolism	Uptake, retention, tissue distribution, amount of parent compound metabolized, metabolites formed and amount, enzyme induction (P-450), biotransformation	Vertebrate	Glickman et al. (1982); McKim and Heath (1983)

TABLE 21. REPRESENTATIVE MAMMALIAN ALTERNATIVE TOXICITY TEST SYSTEMS

Test Type	Test Endpoint	Test Organism	Reference
MUTAGENESIS			
Mammalian cells *in vitro*	Cell toxicity, chromosome breakage, genome mutations sister chromatid exchange	Rodent	Moutschen (1980)
TERATOGENESIS			
Whole rodent embryo culture	Increase in somites, crown-length, protein, DNA, and malformations	Rodent	Fantel (1982)
Embryonic limb bud organ culture	Cell proliferation, differential growth, morphogenetic cell death, chondrogenesis, size and shape of limb parts, collagen and proteoglycan biosynthesis	Rodent	Kochhar (1982)
Palate mesenchyme culture	Morphogenesis of cell	Human	Pratt et al. (1982)
Tumor cell attachment	Number attached to surface	Rodent	Braun et al. (1982)
Prechondrocytes	Colonies formed and staining with Alcian blue	Rodent	Hassell and Horigan (1982)
CARCINOGENESIS			
"Baby rat" liver assay	Development of tumors	Rodent	
Human cell transformation *in vitro*	Transformation of fibroblasts	Human	Borek and Mohr (1982)
REPRODUCTIVE TOXICITY			
Mouse reproduction assay	Hormone receptors, hormone concentrations, morphological evaluation of reproductive tissues, ovulation rates, seminal characteristics, fertilization rates and litter size, spermatogenesis	Rodent	Amann (1982)
NEUROTOXICITY			
Rodent neurobehavioral assay	Sensory loss, muscular weakness, nerve fiber regeneration	Rodent	Tilson et al. (1979) Weiss (1983)
Neuroblastoma and glioma cell cultures	Enzyme activity, differentiation, distribution, viability	Rodent	Prasad (1983), Nardone (1983)

TABLE 21. (Continued)

Test Type	Test Endpoint	Test Organism	Reference
NEUROTOXICITY (Continued)			
Visual-evoked response	Alteration of visual-evoked potential	Rodent	Dyer (1983)
Mouse spinal cord culture	Enzyme activity, differentiation, distribution, viability	Rodent	Nardone (1983)
ORGAN SPECIFIC TOXICITY			
Renal medullary cell culture	Morphological and biochemical alterations	Rodent	Bach (1983)
Rat hepatocyte culture	Morphological and biochemical changes, plasma membrane perturbations	Rodent	Acosta (1983)
ACUTE TOXICITY			
Rabbit corneal epithelial cells	Cell viability, cell proliferation, keratinization, protein synthesis	Rabbit	Chan (1983)
Stratified epidermal cell culture	Cell irritancy	Rodent	Vaughn (1983)
Cultured human endothelial cells	Metabolite release	Human	Newcombe (1983)
Human leukocytes *in vitro* test	Release of histamine	Human	Levy (1983)

TABLE 22. VALIDATED ALTERNATIVE GENETIC TOXICITY TEST SYSTEMS

Test	Toxicological Endpoint	Route of Exposure	Chemical Response	Time to Endpoint	Performance
MICROBIAL					
A. *Salmonella typhimurium* reverse mutation assay (Ames Test)	Reverse gene mutation	Toxicant & bacterial culture added to overlay agar	All major classes except metals & hormones	3 weeks	Detects 85-90% known carcinogens (extensively tested system--over 2600 chemicals tested)
B. *Escherichia coli* Reverse mutation assay	Reverse gene mutation	Toxicant & bacterial culture added to overlay agar	All major classes except metals & hormones	3 weeks	Not as well characterized or as sensitive as *Salmonella* assay; detects only base pair substitutions pair substitutions
C. Gene mutation in *Asperigillus nidulans*	Forward & reverse gene mutation	Toxicant added to cell suspension	?	7 days	Best use is in battery of tests (some results contrast with Ames Test)
D. Gene mutation in *Neurospora crassa*	Forward & reverse gene mutation	Toxicant added to cell suspension	?	7 days	Best use in battery of tests
E. Mitotic gene conversion in *Saccharomyces cerevisiae*	Primary DNA damage	Toxicant added to cell suspension	Aromatic amines, nitrosamines, carbamates, alkalating agents	2-3 weeks	Relatively insensitive to some chemicals
F. Differential Growth Inhibition of Repair Deficient Bacteria	Primary DNA damage	Toxicant added to well in agar; toxicant added to filter disc on surface of agar	Alkylating agents, nitroso compounds, polynuclear aromatics, pesticides	2-4 weeks	Well suited for detecting chemicals causing frame shift mutations
INVERTEBRATE					
A. Sex Linked Recessive Lethal Test in *Drosophila melanogaster*	Point mutations, deletions	Dependent upon compound & potential route of human exposure	Large number chemicals (reference mutagens & pesticides)	72 hours	More sensitive than any other *Drosophila* test in detecting chemical mutagens)
B. Heritable Translocation Test in *Drosophila melanogaster*	Chromosome breakage & eucentric reunion	Orally by injection, feeding, or by exposure to gases	Large number of chemicals (reference mutagens & pesticides)	Not specified	

POTENTIAL, CONSTRAINTS, AND CONCLUSIONS REGARDING THE INCORPORATION OF ALTERNATIVE TECHNOLOGIES

POTENTIAL FOR INCORPORATING ALTERNATIVE TECHNOLOGIES

The use of animals in toxicity testing cannot be eliminated in the foreseeable future. Currently used models have taken years to develop and validate. However, current testing strategies are varied, have somewhat less than precise components, and typically require substantial amounts of time to conduct at significant costs. The continued efforts to develop and validate alternative systems should result in their integration into practical toxicologic methodologies. In the text below, some advantages and limitations of alternatives are reviewed in addition to a feasible approach to apply them.

Advantages of Alternative Test Systems

The primary advantages of incorporating novel test systems using *in vitro*, nonmammalian, or computer techniques concern time and cost savings and ethical considerations. Alternatives may provide the methodology to better understand the targets and mechanisms of toxicity. In some instances, the new test systems may be more efficient and actually replace existing tests. However, the principal role of alternative technologies is likely to screen and select the compounds of highest concern and eliminate unnecessary testing. This should result in significant reductions of animal use.

Limitations of Alternative Test Systems

Prior to the acceptance of alternative systems, a comprehensive knowledge of their limitations must be obtained. The impacts of two primary limitations must be considered, which result from the fact that the test animal is not man. First, dissimilarities exist between the test model and man such as extent and rate of metabolic pathways. The use of *in vitro* techniques suffers from the absence of recruitment that occurs in a non-isolated *in vivo* system. Generally, responses *in vivo* can be dampened or amplified by various recruitment factors such as phagocytizing blood cells. The second limitation concerns a failure of many alternatives to provide comparative information on sensitivity. It is not a requisite that the sensitivity of an alternative be similar to man, but the relationship between the alternative system and man must be known. Under ideal circumstances, the degree of sensitivity and similarity of metabolism would be the chief factors influencing choice of a model system; but this is not always the case (Oser, 1981).

The above mentioned drawbacks are not limited to alternative technologies but are a concern of current toxicity models as well. The classical approach to extrapolation is to (a) observe and compare toxic responses to compounds with large historic data bases in both animals and man, (b) develop theoretical numerical relationships from these data, and on the basis of this information (c) choose good models for the expected endpoints. Clearly, as the data base grows for an alternative test system, the possibility of incorporating it into this approach will also grow.

Feasibility of Applying Alternative Test Systems

Several considerations should be made in an overview of how alternatives can be inserted into testing regimens. The test battery approach of genetic toxicity studies exemplifies one mechanism of applying short-term alternatives in a hazard assessment program. In addition to the use of a simple test battery, another mechanism of applying alternative test systems is in a tiered, decision-tree approach. This approach was outlined for assessing carcinogenicity by Weisburger and Williams (1981). The protocol would call for the initial use of computer models, followed by short-term assays, and then long-term assays and epidemiological data. This test design integrates alternative technologies and could be applied to many toxicologic endpoints.

THE INTERFACE OF A CHANGING SCIENCE AND REGULATORY POLICY

Solutions

A number of solutions can be explored in attempts to reduce animal usage in toxicologic experimentation. Solutions recently presented by Zbinden (1983) include:

- Modifying precision
- Redefining methods
- Redefining test criteria
- Reducing developmental projects
- Reducing routine testing programs
- Replacing existing methods.

All of these ultimately depend upon the prudent and timely blending of scientific breakthroughs with the regulation of testing activities. The recent emphasis on alternative methodologies is leading to an increased pressure on policymakers to develop innovative means to incorporate advances of the science.

Modifying Precision

Steps to modify the required precision in a given test system are already in place for the LD_{50} test (Zbinden and Flury-Roversi, 1981). Reduction in the number of animals used to assess acute toxicity of a chemical substance leads to a less exact but adequate value. Similar steps might be taken, to a lesser degree, in subchronic and chronic studies. Our preoccupation with low-dose estimation has resulted in a situation where precision is strongly emphasized. A solution to this problem might be to relax the requirements for test precision, thereby reducing the number of animals used in testing while emphasizing a different data assessment procedure such as modeling.

Redefining Methods

Another solution, which would require a dedicated effort to interface science and regulation, is a redefinition of methodologies to reduce animal usage. The use of whole animals in toxicity testing is costly. Results are oftentimes unpredictable and difficult to interpret due to biologic variability. On the other hand, the use of such techniques as cell cultures may provide relatively simple results to interpret. The tradeoff is that these results are usually incomplete, difficult to extrapolate, and must be supplemented with further testing. The plausible approach is to redefine methods required for chemical testing. Steps should be taken to establish more flexible protocol requirements which favor animal reduction techniques and do not jeopardize the prediction of toxicity.

Redefining Test Criteria

A review of criteria used to establish the level of tolerance accepted by a given test system can accompany an attempt to redefine methodologies. Criteria establishment is the most critical task performed by the regulatory toxicologist. With intelligent decisions, the changing of test criteria can serve as a solution without compromising the process of risk assessment.

Reducing Developmental Projects

Animal usage will be affected by reduction in the number of government-funded developmental projects such as those designed to develop test systems within the National Toxicology Program. Although such projects should be subject to stringent periodic areview, their elimination is not a reasonable vehicle for reducing animal usage. The potential benefits of these projects underscore their continued existence.

Reducing Routine Testing Programs

A number of routine testing programs exist which may not be justified on a biological basis. In cases where this occurs, a reduction should be considered. Zbinden (1983) has cited the assessment of reproductive toxicity of drugs intended for use only by geriatric populations as an example of such testing programs.

Replacing Existing Methodologies

In some cases, alternative tests or test systems might provide a more efficient means of evaluating a chemical's toxicity. Under these circumstances, a reduction in animal usage may follow by direct replacement with the alternative technology.

Changes Required to Accommodate Alternative Technologies

Many of the changes required to accommodate alternative technologies have already begun to take place. Approaches will become more ordered. Tier testing and test batteries, employing short-term animal and non-animal systems, will facilitate early detection of toxicities with decreasing numbers of animals while providing more knowledge on mechanisms leading to toxicity. This latter phenomenon will indirectly lead to reduced animal use when future compounds of related chemical makeup are evaluated. A more extensive application if *in vitro* systems will be made, as will the use of models to formulate theories and test them.

A critical need currently exists for re-establishing testing regulations, criteria, and methodologies to better reflect the advancing science. Traditional ideas should be thoroughly evaluated and changed to incorporate the philosophy that new approaches, including those which reduce animal usage, be implemented in a timely and effective manner. Current regulations clearly inhibit the development of alternatives. Clear-cut regulatory incentives should be developed which foster the use of logical models and pathways in human safety evaluation.

REFERENCES

Acosta, D., and M. B. Sorensen. 1983. Hepatotoxicity: An *in vitro* approach to the study of toxicity and membrane interactions of cadmium using cultured rat hepatocytes. Johns Hopkins Ctr. for Altern. to Anim. Testing Newsl., Vol. 1 No. 2.

Alder, S., C. Janton, G. Zbinden, eds. 1981. Preclinical safety requirements in 1980. Swiss Federal Institute of Technology and University of Zurich, Switzerland. 89 pp.

Amann, R. P. 1982. Use of animal models for detecting specific alterations in reproduction. Fund. Appl. Toxicol. 2:13-26.

Bach, P. H. 1983. The use of cultured renal medullary cells as an alternative to evaluating chemically related nephrotoxicity in animals. Johns Hopkins Ctr. for Altern. to Anim. Testing Newsl. Vol 1 No. 2.

Bailey, G., M. Taylor, D. Selivonchick, et al. 1982. Mechanisms of dietary modification of aflatoxin B_1 carcinogenesis. Basic Life Sci. 21:149-165.

Balls, M., M. A. Santos, and R. Clothier. 1982. The long-term maintenance of differentiated tissues *in vitro*: Paracetamol metabolism and toxicity in organ culture, pp. 25-47. *In*: Differentiation *In Vitro*: British Society for Cell Biology Symposium 4. M. M. Yeoman, D.E.S. Truman, eds. Cambridge University Press, Cambridge, Great Britain.

Bawden, D. 1983. Structure-toxicity studies, pp. 151-154. *In*: Animals and Alternatives in Toxicity Testing, M. Balls, R. J. Riddell, and A. N. Worden, eds. Academic Press, Inc., New York.

Bondy, S. C. 1982. Neurotransmitter binding interactions as a screen for neurotoxicity, pp. 25-50. *In*: Mechanisms of Actions of Neurotoxic Substances. K. N. Prasad and A. Vrenadakis, eds. Raven Press, New York.

Borek, C., and U. Mohr. 1982. *In vitro* and *in vivo* carcinogenesis: workshop report, pp. 491-493. *In*: Mechanisms of Chemical Carcinogenesis, C. Harris and P. Cerutti, eds. Alan R. Liss, Inc., New York.

Bournias-Vardiabasis, N., R. L. Teplitz, G. F. Chernoff, et al. 1983. Detection of teratogens in the *Drosophila* embryonic cell culture test: Assay of 100 chemicals. Teratology 28:109-122.

Braun, A. G., B. B. Nichinson, and P. B. Horowicz. 1982. Inhibition of tumor cell attachment to concanavalin A-coated surfaces as an assay for teratogenic agents: approaches to validation. Terat. Carcin. Mutagen. 2:343-354.

Bresch, H. 1982. Investigation of the long-term action of xenobiotics on fish with special regard to reproduction. Ecotoxicol. Environ. Saf. 6(1):102-112.

Browne, C. L., and J. N. Dumont. 1979. Toxicity of selenium to developing Xenopus laevis embryos. J. Toxicol. Environ. Health 5:699-709.

Chan, K. Y. 1983. An in vitro alternative to the Draize Test. Johns Hopkins Ctr. for Altern. to Anim. Testing Newsl., Vol. 1 No. 2.

Chan, P. K., G. P. O'Hara, and A. W. Hayes. 1982. Principles and methods for acute and subchronic toxicity, pp. 1-51. In: Principles and Methods of Toxicology, A. W. Hayes, ed. Raven Press, New York.

Chin, J. B., A. M. Sheinin, and A. M. Rauth. 1978. Screening for the mutagenicity of nitro-group containing hypoxic cell radiosensitizers. Mutation Research 58:1-10.

Chou, J. T., and P. C. Jurs. 1979. Computer assisted structure-activity studies of chemical carcinogens. An N-nitroso compound data set. Journal of Medicinal Chemistry 22:792-797.

Clothier, R. H., J. R. Dewar, M. A. Santos, et al. 1981. A comparative study of the deacetylation of paracetamol by urodele and anuran amphibian organ cultures. Xenobiotica 11(3):149-157.

Conquest, P. L., G. Duraud, J. Laillier, et al. 1977. Evaluation of occular irritation in the rabbit: objective versus subjective assessment. Tox. Appl. Pharmacol. 39:129.

Davis, R. E., and S. Benloucif. 1981. Behavioral investigation of neurotoxicity: The effects of colchicine, lumicolchicine and vincristine sulfate on goldfish optic nerve regeneration. Neurotoxicology 2:419-430.

Draize, J. H. 1959. Appraisal of the safety of chemicals in foods, drugs, and cosmetics. In: Association of Food and Drug Officials of the United States, Austin, Texas.

Draize, J. H., G. Woodard, and H. O. Ralvery. 1944. Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J. Pharmacol. and Exp. Therap. 82:377.

Dyer, R. S. 1983. In vitro approaches to neurotoxicity--a physiologist's perspective, pp. 487-497. In: In Vitro Toxicity Testing of Environmental Agents, Part A. A. R. Kolber, T. K. Wong, L. D. Grant et al., eds. Plenum Press, New York.

Enslein, K., and P. N. Craig. 1982. Carcinogenesis: a predictive structure-activity model. J. Toxicol. Environ. Health 10:521-530.

Enslein, K., T. R. Lander, and J. R. Strange. 1983. Teratogenesis: A statistical structure-activity model. Terratogenesis, Carcinogenesis, and Mutagenesis 3:289-309.

Enslein, K., T. R. Lander, M. E. Tomb, et al. 1983. A predictive model for estimatinng rat oral LD_{50} values, pp. 1-123. In: Benchmark Papers in Toxicology, Vol 1, Princeton Scientific Publishers, Inc., Princeton, N. J.

Enslein, K., T. R. Lander, M. E. Tomb, et al. 1983. Mutagenicity (Ames): A structure activity model. Terat. Carcin. Mutagen. 3:503-513.

European Economic Community. 1979. The 6th Amendment and TSCA. Toxic Substances Journal 1 (4):361-382.

Falke, H. E. 1981. Evaluative van klinisch chemische in hematologische parameters in toxicologisch onderzaek (literatueronderzaek), Report No. V. 81.163/200885. CIVO-Institute TNO, Zeist.

Fantel, A. G. 1982. Culture of whole rodent embryos in teratogen screening. Terat. Carcin. Mutagen. 2:231-242.

Fund for the Replacement of Animals in Medical Experiments. 1983. Report of the Frame toxicity committee. Nottingham, England. February, 1983. 40 pp.

Gentile, J. M., and M. J. Plewa. 1983. The maize-microbe bioassay: A unique approach to environmental mutagenesis. pp. 151-165. In: In Vitro Toxicity Testing of Environmental Agents, Part A. A. R. Kolber, T. K. Wong, L. D. Grant, R. S. DeWoskin, and T. J. Hughes, eds. Plenum Press, New York.

Glickman, A. H., S. D. Weitman, and J. J. Lech. 1982. Differential toxicity of trans-permethrin in rainbow trout and mice. I. Role of biotransformation. Toxicol. Appl. Pharmacol. 66:153-161.

Greenwood, P. J. 193. The influence of an oil dispersant Chemserve OSE-DH on the viability of sea urchin gametes. Combined effects of temperature, concentration, and exposure time on fertilization. Aquat. Toxicol. 4:15-29.

Hadden, S. E. 1983. Labeling of chemicals to reduce risk. Law and Contemporary Problems 46:(3):235-266.

Hagström, B. E., and S. Lönning. 1973. The sea urchin egg as a testing object in toxicology. Acta Pharmacol. Toxicol., Suppl. I, 32:1-49.

Hansch, C., B. H. Venger, and A. Panthananickal. 1980. Mutagenicity of substituted (O-phenylenediamine) platinum dichloride in the Ames test. A quantitative structure-activity analysis. Journal of Medicinal Chemistry 23:459-461.

Hara, T. J. 1979. An electrophysiological test for neurotoxicity in fish. Can. Spec. Publ. Fish. Aquatic Sci., Toxic. Tests Freshwater Org. 44:46-56.

Harber, L. C. and R. L. Baer. 1969. Mechanisms of drug photosensitivity reactions. Tox. and Appl. Pharmacol. Suppl. 3:58.

Hart, L. G., J. E. Huff, J. A. Moore, et al. 1983. The National Toxicology Program's research and testing activities, pp. 191-224. In: Hazard Assessment of Chemicals: Current Developments, Vol 2.

Hassell, J. R., and E. A. Horigan. 1982. Chondrogenesis: a model developmental system for measuring teratogenic potential of compounds. Terato. Carcin. Mutagen. 2:325-331.

Hawkes, J. W., and C. M. Stehr. 1982. Cytopathology of the brain and retina of embryonic surf smelt (Hypomesus pretiosus) exposed to crude oil. Environ. Res. 27(1):164-178.

Haynes, L. W. 1980. Synaptic conditioning in the invertebrate nervous system. Comp. Biochem. Physiol. A 66:13-18.

Heywood, R. 1981. Target organ toxicity. Toxicol. Lett. 8:349-358.

Hill, M. W., E. Neale, and A. D. Bangham. 1981. Acute tolerance to the effects of n-butanol and n-hexanol in goldfish. A behavioral study. J. Comp. Physiol. A 142:61-65.

Hose, J. E., and H. W. Puffer. 1983. Cytologic and cytogenetic anomalies induced in purple sea urchin embryos (Strongylocentrotus purpuratus S.) by parental exposure to benzo(a)pyrene. Mar. Biol. Lett. 4:87-95.

Hushon, J. M., A. W. Klein, W.J.M. Strachan, et al. 1983. Use of OECD premarket data in environmental exposure analysis for new chemicals. Chemosphere 12(6):887.

Jelinek, R. 1982. Use of chick embryo in screening for embryotoxicity. Terat. Carcin. Mutagen. 2:255-261.

Johnson, E. M., and B.E.G. Gabel. 1983. An artificial "Embryo" for detection of abnormal developmental biology. Fundam. Appl. Toxicol. 3:243-249.

King, L. A., and A. C. Moffatt. 1981. Hypnotics and sedatives: an index of fatal toxicity. Lancet 2:387-388.

Klaassen, C. D., and J. Doull. 1980. Evaluation of Safety: Toxicologic Evaluation, pp. 11-27. In: Casarett and Doull's Toxicology: The Basic Science of Poisons, J. Doull, C. D. Klaassen, M. D. Amdur, eds. Macmillan Publishing Co., Inc., New York.

Kocan, R. M., M. L. Landolt, J. Bond, et al. 1981. In vitro effect of some mutagens/carcinogens on cultured fish cells. Arch. Environ. Contam. Toxicol. 10:663-671.

Kocan, R. M., M. L. Landolt, and K. M. Sabo. 1979. In vitro toxicity of eight mutagens/carcinogens for three fish cell lines. Bull. Environ. Contam. Toxicol. 23(1-2):269-274.

Kochhar, D. M. 1982. Embryonic limb bud organ culture in assessment of teratogenicity of environmental agents. Terat. Carcin. Mutagen. 2:303-312.

Koenig, C. C., D. C. Abel, C. W. Klingensmith, et al. 1982. Usefulness of the self-fertilizing cyprinodontid fish, *Rivulus marmoratus*, as an experimental animal in studies involving carcinogenesis, teratogenesis and mutagenesis. U.S. Environmental Protection Agency, Environmental Research Laboratory, EPA-600/3-82-075, Gulf Breeze, FL. (NTIS, PB82-249194). 139 p.

Levy, D. A. 1983. An *in vitro* test for irritancy of chemicals using human leukocytes. Johns Hopkins Ctr. for Altern. to Anim. Testing Newsl., Vol. 1 No. 2.

Malling, H. V. 1978. Mutation testing systems, pp. 35-69. *In*: Handbook of Teratology, J. G. Wilson and F. C. Fraser, eds., Vol 4. Plenum Press, New York.

Marion, M., and F. Denizeau. 1983. Rainbow trout and human cells in culture for the evaluation of the toxicity of aquatic pollutants: A study with lead. Aquat. Toxicol. 3:47-60.

McKim, J. M., and E. M. Heath. 1983. Dose determinations for waterborne 2,5,2',5'-[^{14}C] tetrachlorobiphenyl and related pharmacokinetics in two species of trout (*Salmo gairdneri* and *Salvelinus fontinalis*): A mass-balance approach. Toxicolo. Appl. Pharmacol. 68:177-187.

Moutschen, J. 1980. Cytogenetical tests on mammalian cells *in vitro*. Arch. Toxicol. 46:61-75.

Muller, H., and H. P. Kley. 1982. Retrospective study of the reliability of an "approximate LD_{50}" determined with a small number of animals. Arch. Toxicol. 51:189-196.

Nardone, R. M. 1983. Neurotoxicity testing: an *in vitro* strategy. Ann. N.Y. Acad. Sci. 407:458-459.

National Research Council. 1977. Regulations under the Federal Hazardous Substances Act, pp. 119-124. *In*: Principles and Procedures for Evaluating the Toxicity of Household Substances. National Academy of Sciences, Washington, D.C.

National Research Council. 1984. Toxicity Testing: Strategies to Determine Needs and Priorities. National Academy of Sciences, Washington, D.C., 382 pp.

Newcombe, D. 1983. Cultured human endothelial cells: markers for chemical irritants. Johns Hopkins Ctr. for Altern. to Anim. Testing Newsl. Vol. 1 No. 2.

Olson, D. L., and G. M. Christensen. 1980. Effects of water pollutants and other chemicals on fish acetylcholinesterase (*in vitro*). Environ. Res. 21(2):327-335.

Organization for Economic Cooperation and Development. 1981. OECD guidelines for the testing of chemicals. 700 pp.

Organization for Economic Cooperation and Development. 1983. First Addendum to OECD Guidelines for Testing of Chemicals. 41 pp.

Oser, B. L. 1981. The rat as a model for human toxicological evaluation. J. Toxicol. Environ. Health 8:521-542.

Pagano, G., A. Esposito, P. Bove, et al. 1982. Arsenic-induced developmental defects and mitotic abnormalities in sea-urchin development. Mutat. Res. 104:351-354.

Prasad, K. N. 1983. Use of cultures of neuroblastoma and glioma as a model system to study heavy metal-induced neurotoxicity, pp. 421-472. In: In Vitro Toxicity Testing of Environmental Agents, Part A. A. R. Kolber, T. K. Wong, L. D. Grant et al., eds. Plenum Press, New York.

Pratt, R. M., R. I. Grove, and W. D. Willis. 1982. Prescreening for environmental teratogens using cultured mesenchymal cells from the human embryonic palate. Terat. Carcin. Mutagen. 2:313-318.

Pritchard, J. B., and D. S. Miller. 1980. Teleost kidney in evaluation of xenobiotic toxicity and elimination. Fed. Proc. 39(14):3207-3212.

Rose, S. L. and P. C. Jurs. 1982. Computer assisted studies of structure-activity relationships of N-nitroso compounds using pattern recognition. Journal of Medicinal Chemistry 25:769-776.

Sabourin, T. D., and L. B. Goss. 1984. Study of alternative species for biological testing. Final report to the Electric Power Research Institute, Energy Analysis and Environment Division, 141 pp.

Schultz, E. and H. Fuchs. 1982. A new approach to minimizing the number of animals used in acute toxicity testing and optimizing the information of test results. Arch. Toxicol. 51:197-220.

Smith-Sonneborn, J. 1983. Use of a ciliated protozoan as a model system to detect toxic and carcinogenic agents, pp. 113-137. In: In Vitro Toxicity Testing of Environmental Agents, Part A. A. R. Kolber, T. K. Wong, L. D. Grant, R. S. DeWoskin, and T. J. Hughes, eds. Plenum Press, New York.

Srinivasan, A., and P. C. Kesavan. 1979. Reproductive toxicity of caffeine in Musca domestica. J. Toxicol. Environ. Health 5:765-776.

Tattersall, M. L. 1982. Statistics and the LD_{50} study. Arch. Toxicol. Suppl. 5:267-270.

Tichy, M., and J. Trcka. 1983. Contribution of QSAR analysis to data extrapolation between biological objects, pp. 33-41. In: Quantitative Approaches to Drug Design, J. C. Dearden, ed. Amsterdam: Elsevier/North Holland.

Tilson, H. A., P. A. Cabe, and P. S. Spencer. 1979. Acrylamide neurotoxicity in rats: a correlated neurobehavioral and pathological study. Neurotoxicol. 1:89-104.

Trevan, J. W. 1927. The error of determination of toxicity. Proc. Roy. Soc. 101B:483-514.

Tute, M. S. 1983. Mathematical Modelling, pp. 137-151. In: Animals and Alternatives in Toxicity Testing. M. Balls, R. J. Riddell, and A. N. Worden, eds. Academic Press, Inc., New York.

Underbrink, A. G. et al. 1973. pp. 171-207. In: Chemical Mutagens: Principles and Methods for Their Detection, vol. 3. A. Hollander, ed. Plenum Press, New York.

U.S. Congress. 1938. Federal Food, Drug, and Cosmetic Act. 21 USC 301(1976).

U.S. Congress. 1957. Poultry Products Inspection Act. 21 U.S.C. 464.

U.S. Congress. 1967. Federal Meat Inspection Act. 21 U.S.C. 601 et seq.

U.S. Congress. 1970. Clean Air Act. 42 U.S.C. 7401(1977).

U.S. Congress. 1970. Poison Prevention Packaging Act. 15 USC 1471(1982).

U.S. Congress. 1970. Egg Products Inspection Act. 21 U.S.C. 1031.

U.S. Congress. 1970. Occupational Safety and Health Act. 29 U.S.C. 651(1976).

U.S. Congress. 1972. Consumer Product Safety Act. 15 USC 2051(1982).

U.S. Congress. 1973. Lead-based Paint Poisoning Prevention Act. 42 U.S.C. 4801.

U.S. Congress. 1974. Safe Drinking Water Act. 42 USC 300.

U.S. Congress. 1976. Resource Conservation and Recovery Act. 42 USC 6974.

U.S. Congress. 1976. Toxic Substances Control Act. 15 USC 2601.

U.S. Congress. 1966. Fair Packaging and Labeling Act. 15 USC 1451(1976).

U.S. Congress. 1977. Federal Mine Safety and Health Act of 1977. 30 U.S.C. 801 et seq.

U.S. Congress. 1960. Federal Hazardous Substances Act. 15 USC 1261(1982).

U.S. Congress. 1947. Federal Insecticide, Fungicide, and Rodenticide Act. 7 USC 136(1982).

U.S. Congress. Hazardous Material Transportation Act. 49 USC 1803.

U.S. Congress. Port and Waterways Safety Act.

U.S. Congress. 1970. Federal Railway Safety Act. 45 USC 421.

U.S. Congress. Clean Water Act. 33 USC 1251.

U.S. Environmental Protection Agency. 1982. Draft Water Quality Standards Handbook. Office of Water Regulations and Standards, Washington, D.C.

U.S. Environmental Protection Agency. 1978. Director of short term tests for health and ecological effects. EPA Report No. 600/1-78-052. 206 pp.

U.S. Environmental Protection Agency. 1982. Environmental effects test guidelines. EPA Report No. 560/6-82-002. 972 pp.

U.S. Environmental Protection Agency. 1982. Health effects test guidelines. EPA Report No. EPA 560/6-82-001. 432 pp.

U.S. Food and Drug Administration. 1983. Final report on acute studies workshop. Washington, D.C., November 9, 1983. 8 pp.

Vaughan, F. 1983. Irritancy testing with stratified epidermal cell culture. Johns Hopkins Ctr. for Altern. to Anim. Testing Newsl. Vol. 1 No. 2.

Vinson, L. T., U. F. Borselli, W. S. Olentacz, et al. 1969. Laboratory and clinical procedures for assessing photosensitivity potential of tropical agents. Tox. Appl. Pharmacol. Suppl. 3:103.

Weber, L. J., W. H. Gingerich, and K. F. Pfeifer. 1979. Alterations in rainbow trout liver function and body fluids following treatment with carbon tetrachloride or monochlorobenzene, pp. 401-413. In: Pesticide and Xenobiotic Metabolism in Aquatic Organisms. M.A.Q. Khan, I. J. Lech, and J. J. Menn, eds. ACS Symposium Series, No. 99. American Chemical Society, Washington, D.C.

Weis, B. 1983. Behavioral assays in environmental toxicology, pp. 355-381. In: In Vitro Toxicity Testing of Environmental Agents, Part A. A. R. Kolber, T. K. Wong, L. D. Grant et al., eds. Plenum Press, New York.

Weisburger, J. H., and G. M. Williams. 1981. Carcinogen testing: current problems and new approaches. Science: 401-407.

Woodard, G. 1976. Draft survey and evaluation of in vitro toxicity test methods. Draft. U.S. Environmental Protection Agency, Washington, D.C. EPA Report No. 560/5-75-007, 110 pp.

World Health Organization. 1957. General principles governing the use of food additives: First report of the joint FAO/WHO expert committee on food additives. WHO Technical Report Series No. 129. 21 pp.

World Health Organization. 1958. Procedures for the testing of intentional food additives to establish their safety for use: Second report of the Joint FAO/WHO expert committee on food additives. WHO Technical Report No. 144. 19 pp.

World Health Organization. 1967. Procedures for investigating intentional and unintentional food additives: Report of a WHO scientific group. WHO Technical Report Series No. 348. 27 pp.

World Health Organization. 1975. Chemical and biochemical methodology for the assessment of hazards of pesticides for man: Report of a WHO Scientific group. WHO Technical Report Series No. 560. 23 pp.

Zbinden, G. 1983. Konnen tierversuche in der toxikologie eigeschränkt werden, pp. 40-52. In: Bedeutung and Notwendigkeit des Tierversuches in der experimentellen Biologie, P. Walter, ed. USGEB, Basel.

Zbinden, G. and M. Flury-Roversi. 1981. Significance of the LD_{50} test for the toxicological evaluation of chemical substances. Arch. Toxicol. 47:77-99.

Zimmermann, F. K. 1983. Mutagenicity screening with fungal systems. Ann N.Y. Acad. Sci. 407:186-196.

ANIMAL TESTING AND ALTERNATIVES:
REVIEW OF ANIMAL TESTING IN FDA AND EPA; INTERRELATIONSHIP TO PRODUCTS LIABILITY AND; USES OF ALTERNATIVES TO ANIMAL TESTING

BY: MEYER, FALLER AND WEISMAN, P.C.*
4400 Jenifer Street, N.W.
Washington, D.C. 20015

Contract No.: 433-8350.1

---

* David E. Weisman, J.D., L.L.M., Keith A. Rosenberg, J.D., and Bruce M. Mackler, J.D., Ph.D., contributed to this Report. The author also wishes to acknowledge the work of Alan S. Tilles, an Associate of the firm, and the research supplied by Shelley E. Kahn, a Law Clerk with Meyer, Faller and Weisman, P.C.

INTRODUCTION

In recent years, the public has become more conscious of real and potential risks to health posed by chemicals in the environment, in food consumed and in drugs ingested. This heightened awareness has focused attention on the manner in which products and the environment are tested to discover risks and dangers present and how to minimize them.

As a result of the public's growing awareness of real and potential dangers posed to the population from substances in foods, drugs and the environment, science, government and the general public have become more sensitive to the testing of drugs and foods consumed by man and animals, as well as the air, soil and water that comprise the environment and support life itself. An integral part of that testing process has been the use of animals to discover if drugs and foods are safe for use and whether chemicals in the environment are harmful to living creatures.

The Federal government is actively involved in the use of animal testing. Both the Food and Drug Administration and the Environmental Protection Agency promulgate regulations, guidelines and protocols (testing regimes) involving animal testing. The first section of this report concerns animal testing and federal regulations in both the FDA and the EPA. As evidenced from their studies, the testing regimes created are the combined products of science, industry and the Agencies. Exacting regulations for testing are not practical. Rather, less structured guidelines and protocols are more responsive to the flexibility needed to satisfy scientific objectives. Industry balances science against profit, keeping in mind that placing a harmful product into the marketplace can be financially ruinous.

The courts have also become involved in the testing issue because of products liability litigation. In the second section, focus is directed toward products liability law. It seeks to explain the development of this theory of liability as it has evolved from the law of negligence. Of critical importance is the fact that although a manufacturer may satisfy federal testing requirements, that may not be enough to fulfill its legal obligation not to put an unsafe product in the marketplace. The legal duty to test may be greater than that needed to secure federal agency approval. Moreover, there is a continuing duty to test particularly when it becomes known that a product has produced unanticipated side effects. Animal testing has been recognized by the courts as a large factor in proving or disproving the safety of a product in products liability cases.

In the last section, the alternatives for animal testing are explored. Whole animal testing has proven to be both time consuming and expensive. The FDA has been exploring alternatives to determine whether alternative studies are as reliable as animal studies and can be used as a valid substitute therefor.

ANIMAL TESTING AND FEDERAL REGULATIONS IN THE FOOD AND DRUG ADMINISTRATION AND THE ENVIRONMENTAL PROTECTION AGENCY

Introduction:

Pursuant to their statutory authority, the Food and Drug Administration (FDA) and the Environmental Protection Agency (EPA) promulgate regulations concerning animals in testing a variety of substances including food and food additives, drugs, medical devices and chemical substances or mixtures. This section focuses upon the administrative scheme of these agencies as it relates to animal testing summarizing and describing current FDA and EPA regulations, guidelines, protocols and data submission requirements. Attention is also directed to the interrelationship of the Administrative Procedures Act (APA) to the promulgation of regulations relative to animal testing.

I. Overview of Animal Testing and the FDA:

Pursuant to its authority to regulate interstate commerce, Congress has delegated to the Secretary of Health and Human Services and thereafter to the Commissioner of the FDA, broad powers to promulgate regulations consistent with the Agency's statutory objectives. Food, Drug and Cosmetic Act; Federal Caustic Poison Act, 15 U.S.C. §401-411; Fair Packaging and Labelling Act, 15 U.S.C. §1451; Public Health Service Act, 42 U.S.C. §§241-361; 1976 Medical Devices Amendments, 21 U.S.C. §379A. The delegation of rule-making authority to the

Commissioner, FDA, can also be found in various statutory provisions in Food and Drug Administration Enabling Act, 21 U.S.C. §371A; food standards, 21 U.S.C. §341, §347(e); food additives, 21 U.S.C. §348(f); antibiotics, 21 U.S.C. §357F; and new animal drugs, 21 U.S.C. §360B(n)(5).

Through the FDA's broad power to regulate the safety of products, the FDA has created classes of products that require less stringent testing. Previous testing results can be used to demonstrate the safety of these products. Creation of classes of products based upon common characteristics speeds the process by which such products or devices can be approved, lowers the cost of production, and puts the products into the marketplace in a shorter period of time. See, e.g., Drug Efficacy Study Implementation (DESI)[1/], Over The Counter (OTC)[2/] drugs and Generally Recognized as Safe (GRAS)[3/] categories. This regulatory process has received approval by the Supreme Court in Weinberger v. Bentax Pharmaceuticals, Inc., 412 U.S. 645 (1973).

[1/] Drug Efficacy Study Implementation (DESI) - A post-1962 review to evaluate the efficacy of drugs originally approved as safe between 1938-1962, prior to the 1962 Drug Amendments requiring that human drugs be safe and effective.

[2/] Over-the-Counter (OTC) - drugs sold which can be purchased without a doctor's prescription.

[3/] Generally Recognized as Safe (GRAS) - a determination given to a product - drug, food additives, etc., - which, in the opinion of experts, is safe and/or effective based on published or available data.

Through an advisory committee process[4/], outside experts may have an impact upon the decision-making process in creating these categories as well as other testing regimens. While the process has been criticized because of competition among panels as well as the Agency, the advisory committee process plays a prominent role in creating broad common characteristics of products under the FDA's safety and regulatory jurisdiction because of the input from private industry. The advisory committees also play a prominent role in the creation and development of testing requirements.

In addition to regulations, the FDA promulgates guidelines governing principles or practices of general applicability. 21 C.F.R. §10.90. FDA guidelines are not binding, but are advisory in nature. Moreover, guidelines do not include decisions or advice on specific matters. Guidelines are important in the creation of pre-clinical and clinical testing procedures, including scientific protocols. The use of FDA testing guidelines assures acceptance that the procedures are scientifically valid, but does not assure that the product will be approved. The guidelines do, however, express a formal position of the FDA unless amended, and the FDA is obligated to

---

[4/] Advisory Panel - a panel of recognized experts from the industry, medical and/or academic area who serve in an advisory capacity to FDA and make recommendations regarding the interpretation of preclinical/clinical data concerning ingredients and/or products.

follow its own guidelines. Guidelines are public information and can be obtained from the FDA. By regulation, the guidelines are subject to notice and comment provisions and must be made public when published or amended. Guidelines do not establish legal requirements for the safety or validity of products. 21 C.F.R. §10.90.

The broad authority of the FDA to regulate the safety of food, drugs, cosmetics and medical devices, permits the FDA to require protocols clearly indicating the objectives and all methods for the conduct of particular studies. Protocols can apply to pre-clinical, non-clinical and clinical testing as well as laboratory and manufacture practices. For non-clinical laboratory studies that support or are intended to support research or marketing permits for products regulated by FDA, see 42 U.S.C. §262(d); and 21 C.F.R. §58.120; clinical investigation protocols are required for color additive petitions, 21 C.F.R. §71.1, 71.6; food additives, 21 C.F.R. §170.17; medical devices, 21 C.F.R. §812.105; new drugs, 21 C.F.R. §310.3; new drugs for investigational use, 21 C.F.R. §312.1; new drug applications, 21 C.F.R. §314.1, 314.115; OTC drugs, 21 C.F.R. §330.10; new animal drugs for investigative use, 21 C.F.R. §510.3; new animal drug applications, 21 C.F.R. §514; viral vaccines, 21 C.F.R. §630.11; electronic devices, 21 C.F.R. §1010.4. Protocols are the product of either the Agency, the industry or both. They may be amended based upon information obtained during a study.

II. Specific Areas Regulated By FDA That Involve Animal testing

Introduction

Through its regulatory authority, the FDA has developed numerous guidelines and testing protocols relating to animal testing, involving a broad range of products and devices within its jurisdiction. The nature of animals used in testing new drugs or other products or devices varies widely depending upon the type of test being conducted. Although some guidelines are available, the testing protocols are established for each product. These protocols are scientific creations based upon past testing and combine the views of not only the manufacturer, but also those of the Agency and clinicians. Consequently, there are no hard and fast rules governing animal testing. However, some broad conclusions can be drawn from the material below and that is that rodents are generally the preferred animal used in testing, but in tests of longer duration following the use of rodents, other non-rodent animals may be more acceptable.

A. New Human Drugs

The 1938 Amendments to the Federal, Food, Drug and Cosmetic Act required drug manufacturers to submit evidence to FDA that a new drug was safe prior to commercialization. Safety evaluations were primarily based on pre-clinical (animal) studies as well as subsequent patient clinical experiences. The 1962 Amendments required that manufacturers also demonstrate the effectiveness of a new drug; these required either evaluation in appropriate animal models or human clinical experiences.

The FDA has developed over the years various regulatory and informal guidelines as to the nature of the pre-clinical tests recommended to provide sufficient evidence of safety or effectiveness. The general categories of studies include acute toxicity, subchronic toxicity, irritation, pharmacology and metabolic tests, terotogenicity, reproductive, chronic toxicity and carcinogenicity testing. Since each class grouping of drugs, for example anesthetics versus anticancer, have different pharmacologic, toxic, and therapeutic characteristics, the nature of the animal test protocols differ quite widely.

Although the FDA hesitates to define specific pre-clinical protocols for a product, it has developed a series of guidelines which describe the general protocol parameters a manufacturer should consider. Guidelines may be prepared and issued by each drug evaluation group or collectively by the Office of New Drugs. Of the nine guidelines obtained from FDA, the scope, depth and details of each differed widely. For example, two of the guidelines were actually articles in journals, Current Views on Safety Evaluation of Drugs, by Edwin I. Goldenthal (Appendix A), and Guidelines for Reproduction Studies for Safety Evaluation of Drugs for Human Use, by William D'Aquanno (Appendix B). The other seven guidelines appear more formal in nature being documents generated by different drug evaluation groups.

The Guidelines for Pre-Clinical Toxicity Testing of Investigational Drugs for Human Use presents the current FDA recommendations regarding toxicity in laboratory animals. Manufacturers must therefore comply with these generic toxicity tests, choosing which animals are the most appropriate. These guidelines merely specify testing should be done on 3-4 species (one a rodent). The choice of species is left to the manufacturer as the ones most appropriate to the drug. Where evidence exists of potential safety questions, the FDA may request additional tests exceeding the guidelines.

Additionally, manufacturers who have data or sufficient evidence to suggest additional safety concerns have a duty as reasonable manufacturers to perform appropriate animal tests - species, increased numbers, etc. The extent of additional animal testing of new drugs beyond FDA recommendations is dependent on the extent or nature of the information which suggests additional testing may be necessary. Thus, if testing reveals a potential risk to the population, additional testing is required. Moreover, if adverse side effects or increased frequency of expected side effects become known to the manufacturer, more elaborate testing should be done.

The four remaining guidelines are more product oriented - surgical/dental products, anesthetics, anticancer drugs and antiviral drugs. These tend to be in some ways more detailed, but still leave designation of the proper species to the manufacturer. Thus, it is difficult to focus on specific animals since the choice is left up to the individual manufacturer. FDA retains maximum flexibility to request additional species and/or the numbers of animals. This inherent flexibility and the designation of "species" makes any tabulation of guidelines difficult. (See Appendix C, Part I.)

B. Food

Although food safety has been emphasized in the FDA regulatory program since 1906, animal testing of food additives began in the 1950's, initiated by the 1954 Pesticide Chemical Amendments and the 1958 Food Additive Amendments. The Delaney Amendment of 1958 adopted a *per se* ban of any additive which induced cancer from food products. To obtain approval of food additives, manufacturers must demonstrate their substance is "generally recognized as safe" (GRAS) through scientific literature or present data from animal and *in vitro* laboratory tests.

The safety of an additive is generally established by toxicological tests (acute, subchronic and chronic) as well as by carcinogenic, teratogenic, immunotoxic, etc., tests. The regulatory requirements designate an exposure duration of seven to fourteen days, ninety days, twenty-four months, etc., and may specify rodent and non-rodent species (rabbits) depending on the nature of the test. The number of animals is usually left to the manufacturer and probably reflects biostatistical needs given the number of test variables being examined. The guidelines usually recommend the species, number of animals and protocols that should be used for specific products. Again, the wide diversity precludes an effective cataloging of these materials.

Manufacturers frequently exceed the recommended protocols, number of animals, etc., in order to widen the scope of knowledge where potential concerns or future issues are evident. The testing burden is, therefore, on industry to design appropriate protocols. FDA rarely designs appropriate protocols for manufacturers, but rather criticizes the resultant data when it is submitted to FDA in support of an application to market a new product. Manufacturers whose data is rejected because of inadequate protocols must repeat such studies. Thus, the consequences incurred by such manufacturers include delay in obtaining FDA approval to market a product as well as increased expense and usage of laboratory animals. It is, therefore, in the best interest of manufacturers to design statistically appropriate and adequate protocols.

It is important to remember that many animal tests may be performed during the research and development stage prior to the initiation of testing for submission to FDA. While the testing performed for FDA must comply with the Good Laboratory Practices (GLP) regulations, research and development protocols have no such regulatory requirements. Frequently, manufacturers must test many different variables before it can settle on that combination best suited for the intended use. The numbers of animals used, frequency of exposure, duration, scope, etc., of such "pre-GLP compliance" testing are known only to the manufacturer. However, such tests effectively screen compounds or provide a basis for designing appropriate testing protocols which will be submitted to FDA. (*See* Appendix C, Part II.)

C. *Biologic Products*

Biologic products, *e.g.*, vaccines, toxins, diagnostic tests on blood products, plasma products, blood products, etc., are regulated under the Public Health Service Act as well to some extent by the Food, Drug and Cosmetic Act. Thus, the general statutory considerations, with the exception of Section 505 of the Food, Drug and Cosmetic Act, are applicable. The effect is that the same evidentary considerations are required - safety and effectiveness - for FDA approval of new biologics. Investigational biologics are approved pursuant to the same IND format as investigational drugs. The final product license, though analogous to NDA, has its own distinct evaluation scheme.

Biologics, unlike synthesized drugs, are natural substances and have been perceived by the Office of Biologics to require much less animal evaluation as a general rule. However, the Office of Biologics sets standards for each group or class of products, *e.g.*, BCG vaccines, botulism antitoxins, diphtheria antitoxin or vaccine, etc. Manufacturers developing biologics must demonstrate their product(s) meet the standard specifications set for each group of products to obtain approval. Biologics, as compared to drugs, tends to develop more product-oriented regulations guiding the use of animals in products testing. Furthermore, given the proteinaceous nature of the biologics in contrast to chemicals, these safety testing regulations and guidelines tend to be less extensive than those for drugs. Prior to embarking on these regulatory tests, manufacturers have frequently performed extensive "pre-compliance" research testing to establish the safety parameters of their biologics. The FDA required tests, therefore, add to the state of knowledge and are performed in compliance with GLPs; pre-compliance testing frequently does not comply with GLPs and therefore, is unacceptable as evidence to FDA. (*See* Appendix C, Part III.)

D. Medical Devices

The area of medical device testing on animals is much more limited than with drugs. Class I and II devices, as a rule, do not require extensive animal testing. The allegation of "substantial equivalence" to pre-1976 devices are usually based on physical or functional criteria. More recently, 510(k), Premarket Notification applications, contain animal and human test data as another basis to establish "substantial equivalence". Pressure for the inclusion of such data is increasingly asserted by FDA, as substantially equivalent devices become further removed by technological advances from their pre 1976 progenitor device.

The "new" or post '76 devices in Class III more frequently require some animal studies for safety evaluation. This generalization is particularly true for implantable and radiological health devices. The reasonable assurance of Class III device safety required by FDA is quite comparable to that for drugs. However, adequate and well controlled studies with devices in animals, as in humans, are difficult to do. As a result, the guidelines generated are quite product-specific and reflect the general wisdom of the specific device evaluation group and its advisory panel. Animals are frequently not appropriate models for testing the device, thus, the extent of animal testing is quite limited, as compared to drugs. (See Appendix C, Part IV.)

E. Cosmetic Products

The registration of cosmetics with the FDA by manufacturers is on a voluntary basis and since the FDA lacks the statutory power to collect information on products. The FDA also lacks any power to require testing of cosmetics for safety before they are marketed. Manufacturers are, however, subject to misbranding (cosmetic labelling) scrutiny which precludes false or misleading labelling. Marketing a cosmetic which is unsafe and failing to label it as such, would subject the manufacturer to enforcement under the Food, Drug and Cosmetic Act, and possibly the Fair Packaging and Labelling Act.

FDA has, however, developed a series of recommended guidelines for industry use in assessing the safety of cosmetic products. In contrast to other product-oriented protocols, these tend to be quite explicit specifying the type of animal to be used, the duration of exposure, and other testing parameters. Where color additives are involved, the guidelines offer a more general approach. (*See* Appendix C, Part V.)

F. New Animal Drugs, Animal Food Additives and Indirect New Animal Drug Combinations in Non-Food Producing Animals

The federal regulation of animal drugs, medicated feeds, additives, etc., is of rather recent derivation arising from the 1962 and 1968 Drug Amendments. The new animal drug regulations which evolved from the 1968 Animal Drug Amendments were patterned after the 1962 Act. The subsequent 1968 Act focused the statute specifically on animal therapeutics, but kept the definition of safety and effectiveness as the evidentary standard for approval.

The basic intent of these statutes and their resultant regulations is to avoid placing harmful substances (drugs, additives, etc.) in products intended for human food. Secondly, and equally important, is the goal of avoiding toxic, carcinogenic or teratogenic effects in non-food-producing animals, such as pets, and food producing animals. Thus, animal welfare considerations were and remain an essential concern in the implementation of statutes, regulations and guidelines concerning animal drugs, animal feed and additive products.

FDA requires manufacturers to show, prior to commercialization, that new animal drugs are safe and effective through animal toxicity, teratogenic and carcinogenic studies, as well as controlled field studies with the intended species. All such studies must meet minimum and well controlled standards. 21 C.F.R. §514.11. Essentially, the procedures for these investigational studies are quite similar to those for investigational studies of new drugs. However, the requirement that FDA cannot approve any animal product application if the product "induces cancer when ingested by man or animal," imposes a high level of animal testing in appropriate species.

The regulations and guidelines promulgated by the Center for Veterinary Medicine are therefore quite extensive, both as to generic and specific testing recommendations. The diversity of the regulations and guidelines promulgated make classification by animal, number, duration and other convenient classifications impractical. (See Appendix C, Part VIII.)

II. Animal Testing Data in the EPA

Introduction:

In fulfilling its statutory responsibilities, EPA uses toxicity data derived from animal testing in a variety of ways. EPA has the authority to require such data to be submitted under two of the laws it administers, but data are obtained through other means as well. Some data are submitted voluntarily by those who conduct or sponsor testing, some data are obtained from the open literature, and some data are obtained from other government agencies.

This section describes the regulatory programs for which animal testing data are needed, the authorities under which existing data or testing can be required, and the use of testing guidelines in the development animal toxicity data.

A. Overview of EPA Involvement in Animal Testing

1. Office of Air and Radiation

The Clean Air Act (Public Law 95-11, 42 U.S.C. §7401 et seq.), requires the Federal Government and the states to take various actions to improve or maintain the quality of ambient air. Animal testing data supports the following activities under the Act. EPA designates certain substances as "criteria pollutants" and establishes national standards for ambient air based on toxicity and other concerns. EPA also designates certain very toxic pollutants as "hazardous" and establishes standards for their emission or other control.[5/]

The authority for radiation was given to EPA in the President's Reorganization Plan of 1970 (35 Fed. Reg. 15623). Under this plan, EPA makes recommendations to other federal agencies (the Nuclear Regulatory Commission, the Department of Energy and the Occupational Safety and Health Administration) regarding acceptable levels of emissions for tailings and low level wastes. Most of the data used to develop regulatory standards was gathered for humans who were inadvertently exposed to radiation. Animal data is used for genetic effects, dose response relationships and metabolism.

2. Office of Water

The Clean Water Act (Public Law 95-217, 33 U.S.C. §466) requires Federal and state efforts to restore and maintain the integrity of U.S. waters. The Clean Water Act amendments of 1977 contained a list of toxic substances commonly referred to as the

---

[5/] Chemicals are chosen for evaluation based on the likelihood that they are in air. To date, regulated hazardous air pollutants have been supported by human data to justify regulation. Greater reliance on animal data is expected for future regulations. EPA also regulates vehicle emissions, fuels and fuel additives, based in part on toxicity.

126 priority primarily because of their toxic effects on humans and animals. These pollutants are controlled through nationally uniform limitations on the effluents containing these pollutants. Water Quality Criteria have also been promulgated for allowed ambient concentrations of these substances and are used to establish state water quality standards. Now that the work on the priority pollutants is completed, other toxic chemicals are being considered for regulation.

At the present time, animal testing which is of interest under the Clean Water Act is limited to aquatic organisms. Test data on mammals has been considered in the past and may be considered in the future. Such data is presently used in the Drinking Water and Groundwater programs.

The Safe Drinking Water Act (Public Law 95, 42 U.S.C. §300) is designed to protect public drinking water supplies through minimum national standards which are implemented by the states. EPA also regulates the underground injection of fluids and other imminent or substantial hazards to drinking water under this Act. Health advisories are also prepared on specific problems.

Primary Drinking Water Regulations are developed for contaminants which may have adverse effects on human health. Maximum contaminant levels are established using mammalian testing data. Common toxicity tests are used in developing maximum contaminant levels, but acute tests (e.g., LD50) and skin and eye irritations tests are not used because these are not relevant to drinking water problems.

EPA's authority over groundwater is derived from a number of laws. The management of groundwater will be primarily a state responsibility, but EPA will provide technical assistance to state agencies and prepare "advisories" dealing with common problems which endanger groundwater. To some extent, these support activities will rely on toxicity data.

Because groundwater is the source of drinking water for about one-half of the U.S. population, the identification and characterization of groundwater problems is an important part of the drinking water program. Over seven hundred synthetic organic chemicals have been identified in various drinking water supplies. Some epidemiological evidence is available to characterize the toxicity of these contaminants, and more evidence is being collected. Animal testing is also being planned.

3. Office of Solid Waste and Emergency Response

The Resource Conservation and Recovery Act (RCRA)(Public Law 94-580, 48 U.S.C. §6901) protects public health and the environment by controlling the disposal of solid waste and by regulating the management and handling of hazardous waste materials. The EPA is authorized to develop regulations governing the generation, transportation, treatment, storage and disposal of hazardous wastes. These regulations, in addition to state laws on waste, are enforced by the states.

Animal testing is used under §3001, which authorized the development of criteria for the identification of hazardous wastes: toxicity is one of the criteria. RCRA regulations list chemicals which have been determined to be hazardous and processes which are presumed to generate hazardous wastes. Analytical procedures for determining the contents of waste are also described as are criteria for determining whether the contents are toxic (or otherwise hazardous).

Human and animal data are used under RCRA to develop "acceptable daily intakes" (ADIs), used in developing RCRA regulations. Because of the nature of exposure to these wastes, data from short term and dermal tests are not used.

The Comprehensive Environmental Response, Compensation and Liability Act (Public Law 96-510, 42 U.S.C. §9601) authorizes the Federal government to clean up or otherwise respond to the release of hazardous substances or other pollutants which may endanger public welfare. The most significant activity under CERCLA, from the standpoint of animal testing, is the designation of hazardous substances under §102. Substances designated as hazardous under certain sections of other laws (§7 of TSCA, §112 of CAA, §307 of CWA and §3001 of RCRA) are also considered hazardous substances under CERCLA. Under §103, the Administrator is to designate specific amounts of hazardous substances to be "reportable quantities," based in part on toxicity data.

An activity which takes place under §104(i) of CERCLA which diminishes the need for animal testing is the compilation of a Toxic Substances and Disease Registry under the Department of Health and Human Services. This registry will track persons exposed to hazardous substances, along with the medical testing and evaluation which follows the exposure.

4. Office of Pesticides and Toxic Substances

The Federal Fungicide, Insecticide and Rodenticide Act (Public Law 92-516, 7 U.S.C. §136 *et seq.*) is designed to protect human health and the environment from adverse effects of pesticides while allowing the benefits of their use. This is done through the granting of registrations, the approval of labelling and the setting tolerances for maximum residue levels on food which allow safe consumption. EPA also established procedures for the safe storage and disposal of pesticides and pesticide containers.

In registering the approximately 50,000 formulations of "pesticide products," the EPA uses comprehensive registration standards for the approximately 600 "pesticide chemicals" which are the active ingredients. Safety criteria are established for the use of these pesticide chemicals and their formulations. Tolerable exposure levels are determined for consumers of food, field workers, applicators and other persons or organisms that will be unintentionally exposed to pesticides. EPA also approves or disapproves of state registrations for pesticides registered and distribute in only the registering state.

Where pesticide chemicals cause "unreasonable adverse effects," an intensive risk-benefit analysis called Special Review is conducted to determine whether the pesticide should be registered, cancelled, suspended or restricted in its use. Restrictions which are developed are based in part on the mode of action by which the pesticide causes adverse effects (as demonstrated by testing), e.g., inhalation during application, residues on foods and contamination of groundwater.

EPA also relies on toxicity data when it issues emergency exemptions, experimental use permits and temporary tolerances for experimental purposes in response to unexpected and temporary food or health emergencies. Emergency exemptions may be granted to state or Federal agencies for uses not included in the registration. Experimental use permits allow large scale testing of new pesticides or new uses of a registered pesticide.

The Toxic Substances Control Act (Public Law 94-469, 15 U.S.C. §2601) authorizes EPA to regulate chemical substances which present an "unreasonable risk" of injury to health or the environment and to require the reporting or development of data necessary for EPA to assess of risks posed by a given substance.

If a chemical substance presents an unreasonable risk, EPA can regulate its manufacturing, processing, distribution in commerce, use or disposal. Such regulatory actions would be based on toxicity data and exposure data, as well as data regarding the beneficial uses of the substance. Regulation can be in the form of: prohibiting or limiting certain actions; requiring warnings or instructions for use to be given; or, requiring the submission or retention of certain records.

If EPA has reason to believe that a substance presents an unreasonable risk to human health or the environment but does not have sufficient information to make such a finding, it can require reporting of existing toxicity data or exposure data. EPA can also require that a substance be tested for specific toxic effects.

As an aid to identifying chemical substances in need of testing, TSCA required the establishment of the Interagency Testing Committee. This committee suggests chemicals for testing when available information cannot presently support a regulatory decision, but there is cause for concern, i.e., when there is substantial exposure, when a chemical is structurally similar to a known toxin or when the results of screening tests raise safety concerns.

EPA reviews Premanufacture Notifications which must be submitted for "new" chemicals before they can enter commerce. These notifications contain or cite any test data on health or environmental effects known to the submitter, as well as exposure and use information. TSCA distinguishes between "new" and "existing" chemicals based on whether they are on the TSCA Inventory of Chemical Substances.

B. EPA Required Submission of Certain Test Data Under FIFRA and TSCA

1. Federal Insecticide, Fungicide and Rodenticide Act

EPA's Data Requirements for Pesticide Registration specify the kinds of data and information that must be submitted to EPA to support registration of each pesticide under section 3 of the Federal Insecticide, Fungicide and Rodenticide Act. EPA uses the information to determine the identity and composition of pesticides and to evaluate their potential adverse effects and environmental fate.

Tests are either "required" or "conditionally required" depending on factors such as: the results of preliminary tests; whether the pesticide use is for a food crop; whether the use is experimental; where and how the pesticide is to be applied; and, the fate of the pesticide residue.[6/] There is some flexibility in the application of these requirements, but EPA should be consulted if test protocols other than those described are to be used. Flexibility in the data requirements are available through waivers and for minor uses. 40 C.F.R. §158.35, 158.40 (consultation with the Agency; 40 C.F.R. §158.45 (waivers); 40 C.F.R. §158.60 (minor uses); 40 C.F.R. §158.70 (acceptable protocols).

Virtually all data are submitted in the context of obtaining, maintaining or renewing a registration. Under section 6(a)(2) of FIFRA, it is EPA's policy (44 Fed. Reg. 40716 [1979]) that the registrant must also submit any health or safety information regarding a registered pesticide which would be of interest to EPA. This includes the submission of ongoing or completed studies which show previously unknown toxic effects; new toxicity studies for pesticides subject to registration standards, cancellation or review; incidents involving adverse effects to human or nontarget organisms resulting from exposure; or, incidents regarding the lack of efficacy which could indirectly pose a hazard to human life. EPA is presently taking steps to codify and expand these requirements.

[6/] 40 C.E.R. Part 158. The required tests are described in the Pesticide Assessment Guidelines. Although the Guidelines themselves are not part of 40 C.F.R. Part 158, they have undergone public comment and are incorporated by reference.

2. Toxic Substances Control Act

Section 4 is the only section of the Toxic Substances Control Act that gives EPA authority to require testing of industrial chemicals. In requiring testing, EPA must find that:

a) There are insufficient data already available with which to perform a reasonable risk assessment;

b) Testing is necessary to provide such data; and

c) The chemical may present an unreasonable risk of injury to human health or the environment, or the chemical is produced in substantial quantities resulting in significant human exposure or environmental release.

The testing data which would be obtained would then be used in performing risk assessments and in supporting regulatory decisions.

EPA and industry have preferred to negotiate agreements under which industry conducts the tests which may have been required under §4 test rules. It was recently held that these negotiated agreements were not a valid substitute for §4 test rules (cite) that the Interagency Testing Committee recommends. 40 C.F.R. has been set aside for all §4 test rules.

Section 5 requires that before manufacturing or importing of a "new" chemical (a chemical not on the TSCA Inventory of

Chemical Substances) can begin, a Premanufacture Notification must be submitted to the FDA. This is followed by a ninety day Agency review of the submitted data. The submitter must provide all test data in the submitter's possession and control which are related to health or environmental effects, as well as information relating to exposure. If the administrator finds that the information submitted is not sufficient to permit proper evaluation of the health or environmental effects, the chemical can be regulated pending the development and submission of test data. If the submitter withdraws the notification, as often happens, testing would not be required. 40 C.F.R. Part 720.

Section 8 also provides authority for EPA to require submission of certain information in the possession or control of a manufacturer or processor. The general reporting and record-keeping rules promulgated under §8(a) presently apply to about 300 instances. Most of these chemical substances have been designated by the Interagency Testing Committee for possible test rules under §4. Others have been specifically named. Forms are provided for the submission of health and environmental effects data, as well as production and exposure data. The §8(a) reporting requirements can be found at 40 C.F.R. Part 704.

A related provision in §8(d) gives EPA the authority to require manufacturers and processors to submit lists and copies of health and safety studies that they have conducted, that were initiated by them or that were known to them. This information is used in deciding the need for testing under §4 and in identifying ways to reduce risks. The information required under §8(d) is more detailed than that required under §8(a) and need not necessarily be for the same chemicals. 40 C.F.R. Part 716.

Section 8(c) does not involve animal testing, although it can involve animals or affect the need for such testing. This section requires that manufacturers and certain others keep records of any allegation of significant adverse health or environmental effects, including consumer allegations, reports of occupational disease or injury and reports or complaints or injury to the environment. 40 C.F.R. Part 717.

Section 8(e) requires that manufacturers, processors and distributors who obtain information which reasonably supports the conclusion that a chemical substance may present a substantial risk to health or the environment must immediately inform EPA. About one hundred such submissions are received each year. These submissions are evaluated and may lead to §8(a) or §8(d) reporting rules for the substances in question. 40 C.F.R. Part ___. EPA also receives about fifty submissions per year which are not required under §8 but are of interest.

C. Testing Guidelines in EPA

1. Introduction

The purpose of testing guidelines is to describe toxicity tests in such a way that if the guidelines are followed, the results can be assumed to be valid, thus avoiding the possibility of having to repeat a test.

It is necessary for testing guidelines to provide some flexibility. A test substance may not have properties which allow the guidelines to be followed exactly. More importantly, if the guidelines are too strict and become standards, they hinder advances in toxicology.

There are four sets of testing guidelines described in this section: those for the Organization for Economic Cooperation and Development (OECD); those of EPA's Office of Pesticide Programs; those for EPA's Office of Toxic Substances; and, those of the Food and Drug Administration.

2. OECD Guidelines for Testing of Chemicals

The OECD Guidelines for Testing of Chemicals (C.F.R.) have the widest applicability of the four because, under the Mutual Acceptance of Data Decision, the United States and twenty-three other countries agreed to accept data generated according to these guidelines. Guidelines are added or changed by the Updating Panel upon concurrence of member governments. The animal (vertebrate) tests contained in the OECD Guidelines are listed in Appendix D.

B. EPA Guidelines

The Office of Pesticide Programs has developed guidelines for testing required under the Federal Insecticide, Fungicide and Rodenticide Act. These Pesticide Assessment Guidelines contain standards for conducting acceptable tests, guidelines for the evaluation and reporting of data, guidelines regarding when additional testing might be required and examples of acceptable protocols. Similar guidelines have been developed by the Office of Toxic Substances for testing required under the Toxic Substances Control Act.

Unnecessary animal testing is avoided. Both FIFRA and TSCA permit sharing data and the costs of generating it. EPA also encourages those subject to testing requirements to combine two or more tests into a single test wherever possible.

EPA further reduces the use of animals by not requiring testing when the results would clearly not be useful, e.g., if the highest dose level in an acute toxicity test produces no animal deaths, the rest of the test need not be carried out. If a substance has a very high or very low pH, its corrosiveness can be assumed; primary eye and skin irritation tests need not be conducted.

In addition, when a particular chemical is to be tested with the impurities with which it is to be manufactured and sold, EPA allows variations in composition of products for which testing is required if these variations can be expected to have no toxicological significance.

The animal tests involving vertebrates which are contained in the Pesticide Assessment Guidelines are listed in Appendix E. Appendix F contains a similar list of tests developed by the Office of Toxic Substances.

D. Significance of Common Toxicity Tests

1. Acute Toxicity

Acute toxicity tests are used to indicate relative toxicities of substances, the approximate relationship between dose and response and the mode of toxic action. Test substances are administered once orally, dermally or by inhalation. The animals, usually rats, are observed for fourteen days and then sacrificed so that their organs and tissues can be examined for gross changes.

The most common way of expressing the results of an acute toxicity test are the median lethal dose (LD50) or the median lethal concentration (LC50). These terms refer to the dose or concentration at which one-half of the test animals can be expected to die. Acute toxicity tests allow a toxicologist to draw only very general conclusions about the risks to humans. These tests are traditional and inexpensive, and have been carried out for a large number of substances. Thus, they are a convenient tool for identifying and classifying substances of concern.

An abbreviated acute toxicity test is the limit test. A substance is given at a high dose (5 g/kg orally). If none of the animals die from the dose, the full acute toxicity test is normally not required.

2. Short Term Repeated Dose and Subchronic Toxicity

Short term repeated dose and subchronic toxicity test indicate risks to humans in the most common situations, namely, repeated or prolonged exposure. These tests involve repeated exposure to a chemical for fourteen to twenty-eight or ninety days, respectively. Test substances are administered orally, dermally or by inhalation, usually to rats. The term subchronic is used to indicate that the repeated doses occur over less than ten per cent of the animal's lifetime (ninety days).

Effects which are monitored include: accumulation of the test substance in the blood and urine; gross changes in various tissues and organs; and, the presence of infection as an indicator of general health. It is important to evaluate the animals as thoroughly as possible because for many chemicals, these are the only repeated dose studies which will be done.

The reversibility of effects can be determined by the use of a satellite group which is given the highest dose of the test substance for ninety days and then given twenty-eight days to recover before being sacrificed. Satellite groups are generally recommended for subchronic and short term toxicity tests.

C. Chronic Toxicity

Chronic toxicity tests are used to determine the toxic effects of repeated exposure over at least twelve months. These tests detect the effects which require a long latent period to manifest themselves or which are cumulative. These tests are especially important in setting limits for occupational exposure or for widespread environmental contaminants.

Rats and dogs are most commonly used. Exposure can be oral, dermal or by inhalation and continues for twelve and twenty-four months, respectively. Blood and urine are analyzed and tissues and organs are examined for gross and microscopic changes.

D. Carcinogenicity

The objective of carcinogenicity or oncogenicity testing is to observe test animals for cancerous or precancerous tissue. Because of the nature of cancer, testing must be over a major portion of their life span.

Rats and mice are most commonly used and testing is for twenty-four months and eighteen months duration, respectively. Exposure can be oral, dermal or by inhalation. Considerable resources are devoted to the analysis of tissues for abnormal cells once the test is compiled.

E. Combined Chronic Toxicity and Carcinogenicity

Due to the duration of the tests and number of animals needed, chronic toxicity and carcinogenicity testing are often combined, and standard procedures have been developed for combining them.

F. Skin and Eye Irritation

Skin and eye irritation tests are used to develop precautionary information for situations in which these kinds of exposure are possible, i.e., handling of liquids and fine particles. These tests involve a single several hour exposure, usually in rabbits. The observation period can be from three to twenty-one days. The local effects on the skin or the eye are important because this is a common form of exposure in minor accidents.

G. Sensitization

Sensitization tests require an initial exposure and a subsequent exposure within a week to detect whether a hypersensitive state has been established (allergy). Such tests are normally done only for the skin and the eye because these are the most common sites of sensitization.

H. Neurotoxicity

Neurotoxins can damage the nervous system in a variety of ways and several tests have been developed to detect these different kinds of changes, *e.g.*, unsteady gait, diminished reflexes or changes in nerve cell structure.

I. Teratogenicity and Reproduction

Teratogenicity and reproduction tests indicate whether a substance can cause problems in reproduction, such as reduced fertility, birth defects, miscarriage or other toxic effects to the fetus. Exposure can be in males before copulation or in females before or during pregnancy. Effects on subsequent generations can also be determined.

J. Toxicokinetics

Toxicokinetics is the study of the rates of absorption, distribution, metabolism and excretion of toxic substances. Such studies aid in the evaluation of the results of other studies and in the extrapolation of the data to humans. These studies are also helpful in selecting dose levels in other studies.

K. Mutagenicity

Mutagenicity tests can be used to determine if a substance might produce heritable changes in genetic material (DNA) or a developing fetus which could lead to birth defects or carcinogenicity. Some tests for mutagenicity require the use of animals; some do not.

Conclusion

Both the FDA and EPA have substantial regulatory authority to require extensive testing involving the use of animals. In a regulatory sense, much of the testing requirements are left to less rigid guidelines and protocols. Specifics regarding animal testing appear to evolve from prior testing experience and desired testing objectives, combining input from the agencies, manufacturers and the scientific community. This amalgam of past scientific information, product design and objectives of testing is directly related to the types of animals used for specific tests.

It appears that promulgating exacting regulations to test different products or devices for safety or measure environmental toxicity may be impractical. The regulatory process is probably too time-consuming and cumbersome to keep pace with advances in technology and the wave of new products and devices seeking entry in the marketplace. While less structured guidelines and protocols may pit agency against manufacturer through competition in advisory groups where scientific and business interests may clash, the development of testing regimens must have a degree of flexibility in their specific objectives and the animals used to obtain the desired results. In the ultimate, the manufacturer must realize that if it places an unsafe product into the marketplace, or a pollutant or toxin into the earth, air or water, the financial consequences could be devastating. Thus, the incentive for industry to conduct full and adequate testing involving animals cannot be underestimated.

Animal Testing and Product Liability Litigation

This section explores the legal sufficiency of testing methods in relation to the law of products liability. Satisfaction of government testing requirements does not necessarily fulfill the manufacturer's requirement to produce a safe product. Where additional testing would have reasonably revealed side effects, the duty to test may not have been satisfied. Moreover, a manufacturer may have a continuing duty to test when information is available that side effects have resulted once the product has been made available to the public. In sum, satisfaction of government approved or required testing does not translate into protection of the manufacturer in the products liability arena.

Consequently, alternatives to animal testing may satisfy government standards, but not the courts. The interrelationship between an agency-approved testing protocol and its viability as a legal standard must be addressed by manufacturers and the government in adopting statutory and regulatory testing regimens. Legislation and/or regulation may speed the process whereby drugs or devices are approved but may fall short of meeting the legal requirements of tort law.

I. Products Liability In General

Traditionally, persons injured by a product or device were limited by theories of contract or negligence when seeking damages. A principle bar to recover in contract was the inability to demonstrate privity with a manufacturer. Thus, a person who bought a lawnmower and was injured when the blade fell off may be permitted to recover from a manufacturer under a theory of breach of implied warranty, but a neighbor who borrowed the device to mow his lawn could not because he was not the purchaser of the product. In negligence law, liability was limited by notions of "zones of foreseeability," breach of duty and contributory negligence.

The law responded to limitations on liability through the development of strict or products liability, removing traditional barriers to recovery that had been imposed by contract and negligence theories. In large measure, strict liability was a response to sociological and technological changes in the production, distribution and use of products. Technology resulted in mass production of goods transported over great distances in a short space of time. While the speed by which goods were placed into consumption increased, the distance between the manufacturer and the consumer increased because of distribution schemes.. The "middlemen" that separated the consumer from the manufacturer made traditional notions of tort and contract law difficult to apply. The manufacturer no longer had exclusive control over the product from manufacture to user, destroying privity and making negligence more difficult to prove.

Strict liability attempts to respond to these realities. It requires proof that: (1) a product is unreasonably dangerous; (2) the defect existed at the time the product left the control of the manufacturer; (3) that the consumer was injured; and (4) that the product caused the injury. Contributory negligence and lack of privity are not bars to recovery in strict liability.

Recently, Congress has responded to the law of strict liability by tinkering with legislation abolishing strict liability in certain cases. A popular proposal is to establish a single standard of care in products liability law or to replace strict liability with more traditional notions of negligence. While it is not within the scope of this exercise to comment on such legislation, the difficulties inherent in such proposals emerge in the following analyses as strict liability is discussed in terms of drugs and testing procedures.

Generally, a manufacturer has a duty to produce a safe product. "A manufacturer's duty to produce a safe product, with appropriate warnings and instructions where necessary, rests initially on the responsibility each of us bears to exercise care to avoid unreasonable risks of harm to others."[7] This duty extends to all persons who might foreseeably be injured by the product manufactured.[8] This failure to produce a safe product results in liability for the manufacturer for the damages thereby caused.[9]

The manufacturer of a drug which causes injury will be strictly liable in tort for the damages where the foreseeable damage-causing perpensities of the product could have been discovered through the use of testing (animal or otherwise). The product can not reasonably be inspected by the consumer for such defects. Liability will also fall on the manufacturer that fails to adequately warn the consumer about the risks of the drug.[10] Strict liability is liability imposed upon a manufacturer regardless of any showing of negligence on the part of the manufacturer. This strict liability will apply to the manufacturer regardless of whether the manufacturer exercised all reasonable care in production, including the use of animal testing.[11/]

---

[7] Hall v. E.I. Du Pont De Nemours & Co., Inc., 345 F.Supp. 353, 360 (E.D.N.Y. 1972).

[8] Pease v. Sinclair Ref. Co., 104 F.2d 183 (2d Cir. 1939); Lockett v. General Electric Co., 376 F.Supp. 1201, 1207 (E.D.Pa. 1974).

[9] Uniform Commercial Code, Section 2-314(2). According to the Uniform Commercial Code, a safe product is a product which is made in a manner fit for the ordinary purposes for which such goods are sold.

[10] Restatement of Torts, Section 402A; Greenman v. Yuba Power Products, Inc., 59 Cal.2d 57, 377 P.2d 897, 27 Cal.Rptr. 697 (1962); Borel v. Fibreboard Paper Products Corporation, 493 F.2d 1076 (5th Cir. 1973); Lartigue v. R.J. Reynolds Tobacco Co., 317 F.2d 19 (5th Cir. 1963); Toole v. Richardson-Merrell, Inc., 60 Cal.Rptr. 398 (1967).

The number of persons injured by a drug is not necessarily controlling. A manufacturer may be liable even where only a single plaintiff has been injured. Recovery is not precluded until a "statistically significant" number of people have been injured.[12/] The elements of proof are: (1) the drug in question was unreasonably dangerous; (2) the defect existed at the time the drug left the manufacturer's hands; (3) that the consumer was injured or suffered damages from the use of the drug; and (4) that the defect in the drug was the proximate cause of the injury.[13/] It is important to note that a drug may be unreasonably dangerous even though it is manufactured exactly as it was intended by the manufacturer (i.e., not defective).[14/] An unreasonably dangerous drug is one which is so dangerous that a reasonably prudent manufacturer would not sell the product if it

---

[11/] *Reyes v. Wyeth Laboratories*, 498 F.2d 1264 (5th Cir. 1974); *Todd Shipyards v. Turbine Serv., Inc.*, 467 F.Supp. 1257 (E.D.La. 1978).

[12/] *Ferebee v. Chevron Chemical Co.*, 736 F.2d 1529, 1536 (D.C. Cir. 1984).

[13/] *Id.*; *Nicklaus v. Hughes Tool Company*, 417 F.2d 983 (8th Cir. 1969).

[14/] *Reyes*, *supra*.

knew the risk involved.[15/] In reaching this determination as to whether the drug is unreasonably dangerous even though it is manufactured exactly as intended, it is necessary to weigh the dangers of the harm of the drug versus the utility of the good which the drug may bring. It is important to note that drugs may also be unreasonably dangerous as marketed. This was a result from a drug not being accompanied by the proper directions and warnings. The failure to warn is a "defect". [16/]

II. The Duty to Warn

Not only does a manufacturer have a duty to produce a safe product, the manufacturer also has a duty to warn of all dangers discovered by its animal or other testing.[17/] For prescription drugs, this duty to warn is generally limited to the duty to warn the physician involved, except where the prescription drug will not be dispensed by a physician.[18/] An exception to this is the

---

[15/] Borel v. Fibreboard Paper Products Corp., 493 F.2d 1076 (5th Cir. 1973); Ferrigno v. Eli Lilly and Co., 175 N.J.Super. 420, 420 A.2d 1305 (1980); Reyes, supra.

[16/] Chambers v. G.D. Searle & Co., 441 F.Supp. 377 (D.Md. 1975); Borel, supra; Reyes, supra.

[17/] Ferrigno, supra; O'Hare v. Merck & Company, 381 F.2d 286 (8th Cir. 1967).

[18/] Reyes, supra; Stottermire v. Cawood, 215 F.Supp. 266 (D.D.C. 1963); Brick v. Barnes-Hines Pharmaceutical Company, Inc., 428 F.Supp. 496 (D.D.C. 1977); Smith v. E.R. Squibb & Sons, Inc., 245 N.W.2d 52 (Mich.App. 1976); Reyes, supra; Mulder v. Parke-Davis Company, 416 F.2d 417 (2nd Cir. 1969); Hoffman v. Sterling Drug,
(Footnote Continued)

case of birth control pills. At least one court has found it important to provide warnings not only to the physician but also the consumer in order to adequately inform patients of the possible side effects of the drugs and to enable the patient to identify symptoms which might otherwise go undetected.[19] In other cases warnings may be inadequate, such as where a dangerous product is likely to be used or misused by children, especially if the danger can be reduced by a different design or by other manufacturing methods. The duty to warn also arises when a departure from the directions provided by the manufacturer may create a serious danger. However, the manufacturer is generally under no duty to warn against obvious hazards, or abnormal, unforeseeable uses.[20] The manufacturer must warn those whom it reasonably expects to use the drug.[21] The warning provided by the manufacturer may be weakened by a definitive statement of

---

(Footnote Continued)
Inc., 485 F.2d 132 (3rd Cir. 1973); Timm v. Upjohn Company, 624 F.2d 536 (5th Cir. 1980); Davis v. Wyeth Laboratories, 399 F.2d 121 (9th Cir. 1968); McCue v. Norwich Pharmacal Company, 453 F.2d 1033 (1st Cir. 1972).

[19] Lukaszewicz v. Ortho Pharmaceutical Company, 510 F.Supp. 961 (E.D.Wisc. 1981); Contra; Chambers, supra; Dunkin v. Syntex Laboratories, Inc., 443 F.Supp. 121 (W.D.Tenn. 1977); Leibowitz v. Ortho Pharmaceutical Co., 224 Pa.Super. 418, 307 A.2d 449 (1973).

[20] Jamieson v. Woodward & Lothrop, 101 U.S.App. D.C. 32, 247 F.2d 23, cert.denied, 355 U.S. 855 (1975); Restatement 2d, Torts §395, Comment j.

[21] Weekes v. Michigan Chrome & Chem. Co., 352 F.2d 603 (6th Cir. 1965).

safety, thereby exposing the manufacturer to liability.[22/] An adequate warning, however, will generally relieve a manufacturer from liability.[23/] The manufacturer's duty to warn continues after the sale of the product. Therefore, a manufacturer must attempt to warn consumers who have purchased the drug if the manufacturer discovers a defect or other type of problem in subsequent testing or doctors' reports after sales have taken place.[24/] This duty to warn extends to a duty to warn those allergic users where the class of allergic users is potentially appreciable or the harm which may be caused is great.[25/]

III. Manufacturer's Liability For Low-Utility Drugs

A manufacturer may be strictly liable where it produced a prescription drug which could not reasonably have appeared to be useful and desirable at the time of its manufacturer, or where, despite some apparent efficacy, the medically recognizable risk foreseeable in the drug outweighed its utility.[26/] Factors which

---

[22/] Maize v. Atlantic Ref. Co., 352 Pa. 51, 41 A.2d 850 (1945).

[23/] Chambers, supra; Dunkin, supra; Ferrigno v. Eli Lilly and Co., 175 N.J.Super. 420, 420 A.2d 1305 (1980); Leibowitz, supra.

[24/] Ferrigno, supra; Leibowitz, supra; Schenebeck v. Sterling Drug, Inc., 423 F.2d 919 (8th Cir. 1970).

[25/] Restatement Torts, Section 402A, Comment j.

[26/] Ferrigno, supra.

a court would consider in reaching a determination as to a drug's utility include: (1) overall usefulness and benefit to the user and the public as a whole; (2) likelihood that the product would cause injury and the seriousness of such injury; (3) availability of a substitute product at the time of sale and distribution which would meet the same need as the product, but would not be as unsafe; (4) ability of the manufacturer to eliminate the unsafe character of the product without impairing its usefulness; and, (5) the expense involved to the manufacturer in eliminating the unsafe character of the product.[27/] In other words, a higher standard of care will apply when there is an alternative drug already available, which is effective and does not carry with it some other highly undesirable side effects.

The most striking example of this is the case of Quadrigen. Quadrigen was a drug created by Parke-Davis and Company which contained four (4) antigens; diphtheria toxoid, tetanus toxoid, pertussis (whooping cough) vaccine, and poliomyelitis vaccine. The purpose for the creation of the drug was to combine the company's trivalent antigen sold under the trade name "Triogen", which contained diphtheria toxoid, tetanus toxoid and pertussis vaccine, with a poliomyelitis vaccine in order to eliminate one of the two vaccines which infants usually receive. Because of the use of a particular preservative in the drug, Quadrigen was

---

[27/] Ferrigno, supra at p. 1319.

the proximate cause of encephalitis in some infants. Parke-Davis was found strictly liable in at least two cases.[28/] There was no tremendous public need for the new drug, since the old Triogen and separate polio vaccine were adequate, although there was some desirability to have one shot instead of two.[29/]

IV. Exceptions To Strict Liability: Unavoidably Unsafe Drugs And Unforeseeable Side Effects

There are several exceptions to strict liability for manufacturers of drugs. Most of these exceptions are covered by the Second Restatement of Torts, Section 402A, comment k. This section deals with those products which are "unavoidably unsafe." It exempts from strict liability those products which, in the present state of human knowledge, are quite incapable of being made safe for their intended and ordinary use. This includes drugs which have a high risk of harmful side effects but treat symptoms which pose a very high danger if left untreated (*i.e.*, rabies vaccines).[30/]

[28/] *Tinnerholm v. Parke-Davis and Company*, 411 F.2d 48 (2nd Cir. 1969); *Stromsodt v. Parke-Davis and Company*, 411 F.2d 1390 (8th Cir. 1969).

[29/] As will be discussed later, Parke-Davis was also found guilty of negligence in manufacturing, testing and failing to give adequate warnings.

[30/] Medical devices are considered to be a drug for purposes of this section. *U.S. v. An Article of Drug ... Bacto Unidisk*, 394 U.S. 784, 792 (1968).

Another exception to strict liability for manufacturers of drugs covers the harmful effects of drugs which no developed skill or foresight could avoid, despite complete animal or other testing performed with state-of-the-art equipment at the time.[31/] This principle of unavoidably unsafe applies throughout the law of products liability. However, as discussed previously, there is a continuing duty on the part of a manufacturer to be aware of the state-of-the-art.[32/] In this regard, for example, a manufacturer cannot ignore reports of ill effects from drugs it has manufactured or similar drugs by other manufacturers after the product has been placed on the market.[33/]

V. Manufacturer's Negligence And The Need To Perform Tests

In those situations where strict liability does not apply (because of one of the previously mentioned exceptions), courts then undertake an examination as to whether the manufacturer was negligent. Basically, a manufacturer's negligence would arise in two situations. First, as mentioned previously, negligence may

---

[31/] Ferrigno v. Eli Lilly and Company, 175 N.J.Super. 420, 420 A. 2d 1305 (1980); O'Hare, supra; Stromsodt v. Parke-Davis and Company, 257 F.Supp. 991 (D.N.D. 1966), affirmed, 411 F.2d 1390 (8th Cir. 1969).

[32/] Schenebeck v. Sterling Drug, Inc., 423 F.2d 919 (8th Cir. 1970); Ferrigno supra; Leibowitz, supra.

[33/] Ferrigno, supra; Tinnerholm v. Parke-Davis and Company, 285 F.Supp. 432 (S.D.N.Y. 1968), affirmed, 411 F.2d 48 (2nd Cir. 1969).

arise in a manufacturer's failure to adequately warn the consumer. The second type of negligence is a manufacturer's negligent testing of a product. In this situation, a manufacturer would be liable for negligent failure to conduct tests (animal or otherwise) of a product which are commensurate with the foreseeable dangers which result as a consequence of the use of the product.[34/] Thus, a greater risk of serious harm from use of a product would mandate a greater amount of testing (animal or otherwise). A manufacturer's testing may be deficient because of methodological weakness in the test conducted or because of a failure to conduct other tests.[35/] The duty on the part of the manufacturer to test a product continues after the manufacture of the product.[36/]

Reasonable foreseeability is a legal term which is applied to general principles of tort law. What is reasonable turns on the facts of a particular case. Testing is an integral part in determining what is reasonably foreseeable. The length and scope

---

[34/] Lockett v. General Electric Company, 376 F. Supp.1201 (E.D.Pa. 1974); Borel, supra; Nicklaus v. Hughes Tool Company, 417 F. 2d 983 (8th Cir. 1969); Hall v. E.I. DuPont De Nemours and Company, 345 F.Supp. 353 (E.D.N.Y. 1972); Todd Shipyards Corporation v. Turbine Service, Incorporated, 467 F.Supp. 1257 (E.D.La. 1978); Fernandez v. Chios Shipping Company, Limited, 542 F.2d 145 (2nd Cir. 1976); O'Hare v. Merck & Co., Inc., 381 S.2d 286 (8th Cir. 1967).

[35/] Anglo Eastern Bulkships Ltd. v. Ameron, Inc., 556 F.Supp. 1198 (S.D.N.Y. 1982).

[36/] Leibowitz v. Ortho Pharmaceutical Corporation, 224 Pa.Super. 418, 307 A.2d 449 (1973).

of the testing impact on the manufacturer's knowledge of potential risks. The manufacturer's duty to conduct tests is dependant upon the foreseeable risk of harm to potential users in light of the current scientific or medical knowledge and discoveries. Whether the manufacturer "was in a position to foresee the risk turns on whether 'the manufacturer knew, or through the exercise of reasonable care should have known, of the existence of the danger.'"[37/] The best example of this are those cases dealing with cigarette smoking. In at least three separate cases, courts have held that a manufacturer of cigarettes can be held liable for foreseeable harm (causing cancer) resulting from a defective condition in cigarettes when the consumer uses cigarettes for purposes for which it was manufactured and marketed. However, there is no absolute liability on the part of the manufacturer for harmful effects of which no developed skill or foresight can avoid. Therefore, cigarette smokers who started smoking before the great cancer causing debate cannot rely on tobacco companies' "warranty" that their cigarettes had no carginogenic element; the manufacturer is not an insurer against the unknowable.[38/]

[37/] Hall, supra at 362 (citations omitted).

[38/] Pritchard v. Liggett and Meyers Tobacco Company, 295 F.2d 292 (3rd Cir. 1961); Ross v. Phillip Morris and Company, 328 F.2d 3 (8th Cir. 1964); Lartique v. R. J. Reynolds Tobacco Company, 317 F.2d 19 (5th Cir. 1963).

Another example of the application of the concept of foreseeable risk of harm is the class of cases dealing with manufacturers of the drug DES. In at least two suits brought by offspring of mothers who used DES during pregnancy, courts have held that manufacturers would be liable for cancer caused by ingesting the drug if such cancer was foreseeable, and the manufacturer failed to test for it or warn of the dangers.[39/] In one case, the New Jersey Supreme Court applied the general rule that there could be no assumption that a manufacturer knew of the harmful character of the drug, if that fact was unknowable at the time of manufacture. However, the court found that the manufacturer was charged with knowledge which could have been acquired through the exercise of reasonable care.[40/] Such reasonable care would include adequate testing which was not performed on pregnant mice. The New York Supreme Court found that cervical cancer in offspring was a foreseeable harm which could come from ingesting DES, and found that the manufacturer was negligent in failing to test for DES's cancer-producing capacity in pregnant mice, once it was decided by the manufacturers to market the product to pregnant women.[41/] It is important to note that a manufacturer may not ignore subsequent developments in the "state-of-the-art", but must continue to

---

[39/] Ferrigno, supra; Bichler v. Eli Lilly and Company, 436 N.Y.S. 2d 625 (1981).

[40/] Ferrigno, supra.

[41/] Bichler, supra.

test, as mentioned previously after the drug has been approved and marketed.[42/]

VI. Adequacy of Manufacturer's Testing

Having established that it is the duty of a manufacturer of a drug to test its product, and to conduct tests relative to the potential danger of the drug, it is next necessary to examine what constitutes adequate testing. Initially, it is clear that a manufacturer must use the safest and most effective inspection or testing method available to fulfill a manufacturer's obligation to conduct tests on animals where an animal test would constitute the most sufficiently recognized testing. For example, it has been established that testing on monkeys is the most reliable test for strains of Polio vaccine.[43/] The requirement to test Polio vaccine on monkeys has even been codified.[44/]

However, there is no duty on the part of the manufacturer to conduct tests which are impracticable or economically unfeasible in relation to the risk of harm from use of the product.[45/]

[42/] Ferrigno, supra; Leibowitz, supra; Schenebeck, supra.

[43/] Griffin v. U.S. 351 F.Supp. 10 (E.D.Pa. 1972).

[44/] 42 C.F.R. Section 73.114(b)(1).

[45/] Livesley v. Continental Motors Corporation, 331 Mich. 434, 49 N.W.2d 365 (1951); Sieracki v. Seas Shipping Company, 57 F.Supp. 724 (E.D.Pa. 1944).

Nevertheless, it is important to point out that the lack of obligation to conduct impracticable or economically unfeasible tests has only been applied in cases of manufacture of non-drug items.

For a plaintiff to prove that a manufacturer was negligent in failing to adequately test a product, he or she must show that the tests conducted were so inadequate as to be unreasonable during the time that the product was tested.[46/] In order to accomplish this the plaintiff need show that the defendant manufacturer failed to exercise reasonable care in choosing testing methods. Further, the plaintiff must suggest other, more suitable tests. For example, in a DES case, the plaintiff showed that the manufacturer failed to perform appropriate tests for DES's foreseeable cancer-producing propensity in pregnant mice.[47/]

There are no hard standards by which adequacy of testing may satisfy the requirements of the law. The question of the adequacy of tests conducted by the manufacturer is to be determined by the judge or jury.[48/] The case law acknowledges

---

[46/] Anglo Eastern Bulkships, Limited v. Ameron, Inc., 556 F.Supp. 1198 (S.D.N.Y. 1982).

[47/] Bichler, supra.

[48/] International Harvester Company v. Sharoff, 202 F.2d 52 (10th Cir. 1953); Pritchard, supra; Bichler, supra.

that animal testing is the most effective method for testing drugs. The factual issue at trial generally becomes whether the manufacturer satisfied the duty to test before and after the drug was marketed.

Compliance with a federal government agency standards for new drug marketing is evidence of exercise of care, but not controlling, as discussed in other sections of this report, testing regimens are the product of science, the Agency and the manufacturer. FDA guidelines specifically are designed to prove the safety and effectiveness of a drug.[49/] In order for a government agency (i.e., FDA) to approve a new drug, a manufacturer must conduct "adequate and well-controlled studies."[50/] But guidelines do not necessarily satisfy legal requirements.[51/]

While the trier of fact may consider whether a manufacturer has complied with agency testing standards, or even complied with industry animal testing standards, the manufacturer's compliance with these standards is not conclusive proof of safety. A

[49/] U.S. v. Articles of Food and Drug, etc., 518 F.2d 743 (5th Cir. 1975).

[50/] U.S. v. Articles, supra; U.S. v. An Article of Drug Neo-Terramycin, 540 F.Supp. 363 (N.D.Tex. 1982).

[51/] Toole v. Richardson, Merrill, Inc., 60 Cal.Rptr. 398 (1967). For the purpose of judicial review under APA, whether the studies are adequate, is a matter which is committed to the expertise of the Agency. In general, a court will not disturb the Agency's determination.

failure on a part of a manufacturer to comply with FDA, other federal government standards or industry animal testing standards, demonstrates negligence,[52/] but compliance with government or industry standards, while evidence of lack of negligence, is not conclusive upon the issue.[53/] Therefore, a jury may find a drug manufacturer negligent for failing to conduct adequate animal tests, despite the fact that the manufacturer complied with all applicable government and industry standards.

In seeking its FDA approval, a manufacturer will conduct initial studies. As discussed in relation to the Polio vaccine, in many situations, a statute will require that animal studies be performed.[54/] Courts have found that a drug manufacturer must conduct both in vivo and in vitro studies in order to obtain FDA approval.[55/]

[52/] Frankel v. Styer, 386 F.2d 151 (3rd Cir. 1967); McComish v. DeSoi, 42 N.J. 274, 200 A.2d 116 (1964); Moren v. Samuel M. Langston Company, 96 Ill.App.2d 133, 237 N.E.2d 759 (1968); Orthopedic Equipment Company v. Eutsler, 276 F.2d 455 (4th Cir. 1960).

[53/] Ferrigno, supra; C.D. Herme, Inc., v. R.C. Wtay Company, 294 S.W. 2d 534 (Ky.App. 1956); Sherman v. M. Lowenstein and Sons, Inc., 28 A.D.2d 922, 282 N.Y.S.2d 142 (1967); Ferebee v. Chevron Chemical Co., 736 F.2d 1529 (D.C. Cir. 1984).

[54/] Griffin, supra.

[55/] U.S. v. An Art. of Drug Con. of 4,680 Pails, 725 F.2d 976 (5th Cir. 1984). A single in vitro study would be insufficient to prove a drug's safety and effectiveness.

A manufacturer seeking FDA approval for a new drug must also demonstrate that the studies conducted were "well-controlled". Several factors enter into this consideration: (1) Whether the test group of animals was large enough;[56/] (2) Whether the drug tested against the drug formerly used to see if the new drug is more effective than the old;[57/] (3) Whether the manufacturer conducted both short term and long term animal and human tests;[58/] and (4) Whether the drug tested as it will be used (strength, temperature, etc.).[59/]

Once a manufacturer has concluded its initial studies, it must then perform clinical studies on humans.[60/] Where applicable, the manufacturer should use double blinding. In addition, these clinical studies should be long term (at least three years).[61/] The only exception to the need for a manufacturer to perform adequate and well-controlled studies is the situation where a manufacturer of a generic drug may rely on the studies of others if the drug is in the exact same form, dosage and application as a previously approved drug presently on

---

[56/] Edison Pharmaceutical v. FDA, 600 F.2d 831 (D.C. Cir. 1979).

[57/] Tinnerholm, supra; Stromsodt, supra.

[58/] Toole, supra.

[59/] Tinnerholm, supra; Stromsodt, supra.

[60/] Edison, supra.

[61/] Toole, supra.

the market.[62/] In that situation the manufacturer need only submit bioavailability studies.[63/]

To summarize, a manufacturer must conduct adequate tests for all foreseeable side-effects. The tests that the manufacturer conducts must be by the safest and most effective method available. Animal testing has generally been recognized as meeting that requirement. While there are no cases directly on point, it would appear that a manufacturer must conduct the safest and most reasonable testing available, regardless of the agency's requirements. This situation may be analogized to that which incurred in Sieracki v. Seas Shipping Company, Inc. In Sieracki, testing of a boom which was to be used on a ship was the subject of litigation. The defendant manufacturer performed tests of the boom, but failed to perform a test which would have revealed the crack which was in the boom. The alternative test was feasible and was not unduly costly, in light of the possible foreseeable harm which could come from a crack in the boom.[64/]

The adequacy of the tests by the manufacturer is a fact question. It turns on the gravity of the possible harm, the utility of the product (i.e., the "Quadrigen case"), and the cost

---

[62/] U.S. v. Undetermined Quant of Various Art., etc., 675 F.2d 994 (8th Cir. 1982).

[63/] Id.

[64/] 57 F.Supp. 724 (E.D.Pa. 1944).

of performing the tests. An example of where adequate testing was performed and proper warning given is the case of contraceptives. Several courts have found that manufacturers marketing contraceptives, having discovered small incidences of stroke and thrombophelitis through testing, have adequately warned of such dangers, given the existing state of knowledge at the time.[65/]

While there are no reported cases where a manufacturer substituted a non-animal test for an animal test or where the manufacturer substituted one type of animal for another, it is nonetheless clear, that where certain animals are the best indicator of toxicity, a manufacturer must use that animal. The tests which the manufacturer conducts must be both long and short term. In addition, the drug must be tested in the form that it will be marketed (strength, chemical content, etc.). This would occur where the plaintiff could show another test on an animal or another test on a different animal which was more suitable, or where the defendant failed to test for a foreseeable side-effect, whether or not that test was required by the government.

While a search of over 500 products liability cases by the authors revealed no case directly on point for manufacturer substitution of testing, there are several examples where

---

[65/] Dunkin, supra; Lukaszewicz, supra; Chambers, supra; Leibowitz, supra.

manufacturers have conducted tests which were inadequate, both drug and non-drug products.

In the "Quadrigen" cases, discussed previously, a manufacturer developed a drug for which other alternatives were available, thus the drug had very low utility. Because of this low utility, the standard of testing was held to be higher. Since there was no great need for the drug, the importance of the availability of the drug was not great, no harm would result because of the drug's non-availability. Clearly, this was not a case where a drug was "unavoidably unsafe." The manufacturer, in testing the drug, failed to account for variations in temperature which would occur to any vaccine being used in the market. Only refrigerated samples of the drug were tested. Because of this, the manufacturer failed to discover the drugs lack of potency at room temperature. This, combined with inadequate controls over clinical studies and the failure to be cognizant of reports in the use of the preservative used in the drug, caused the manufacturer to place on the market a drug which caused encephalitis to some infants. The injury was a reasonably foreseeable consequence of improper and incomplete testing. The potency test, performed by injecting groups of mice with varying dilutions of the vaccine, failed to show the defect of the product, constituting a negligent failure to adequately test. Parke-Davis had met all applicable Government standards. However, the Court found the standards promulgated were minimal.[66/] It was held that Parke-Davis "knew or should have known that Quadrigen might cause encephalopathies in some users and to warn of the danger."[67/]

Another inadequate testing case involved the drug "Aralen". In that situation, the manufacturer conducted three tests on animals which were long-term (i.e., extended longer than one year). The first was a two-year test on rats. The second study lasted seventy-four (74) weeks and was conducted on monkeys. The last test was not conducted until after the complained of effect was well known to the manufacturer. The defendant also conducted twelve to fifteen short terms tests. However, in each of these situations the manufacturer failed to conduct any test specifically directed at any side effects causing harm to a patient's eye or retina. The side effects were held to be foreseeable. In a series of lawsuits, the manufacturer was found liable for failing to test its product adequately for foreseeable dangers and for failure to adequately warn the consumer of possible side effects.[68/] Perhaps the most appropriate rule for

---

[66/] Stromsodt, supra at p. 997.

[67/] Id.

[68/] Hoffman v. Sterling Drug, Incorporated, 485 F.2d 132 (1973); Schenebeck, supra; Basko v. Sterling Drug, Incorporated, 416 F.2d 417 (2d Cir. 1969); Kershaw v. Sterling Drug, Incorporated, 415 F.2d 1009 (5th Cir. 1969); Sterling Drug, Incorporated v. Yarrow, 408 F.2d 978 (8th Cir. 1969); Sterling Drug, Incorporated v. Cornish, 370 F.2d 82 (8th Cir. 1966); Bine v. Sterling Drug, Incorporated, 422 S.W.2d 623 (D.Mo.1968); Krug v. Sterling Drug,

(Footnote Continued)

this area of law is that stated in one of the "Aralen" cases. The court quoted Justice Jackson, dissenting in the case of Dalehite v. United States.[69/] There, Justice Jackson opined:

> This is a day of synthetic living, when to an ever-increasing extent our population is dependant upon mass producers for its food and drink, its cures and complexions, its apparel and gadgets. These no longer are natural or simple products but complex ones whose compositions and qualities are often secret. Such a dependent society must exact greater care than in more simple days and must require from manufacturers or producers increased integrity and caution as the only protection of its safety and well-being. Purchasers cannot try out drugs to determine whether they kill or cure...where experiment or research is necessary to determine the presence or the degree of danger, the product must not be tried out on the public, nor must the public be expected to possess the facilities or the technical knowledge to learn for itself of inherent but latent dangers. The claim that a hazard was not foreseen is not available to one who did not use foresight appropriate to his enterprise.[70/]

(Footnote Continued)
Incorporated, 416 S.W.2d 143 (D.Mo. 1967); Oppenheimer v. Sterling Drug, Incorporated, 7 OhioApp.2d 103, 219N.E. 2d 54 (1964); Cochran v. Brooke, 243 Or. 89, 409 P.2d 904 (1966).

[69/] 346 U.S. 15, 73 S.Ct. 956, 97 L.Ed.2d 1427 (1953).

[70/] Id. at 51-52, quoted in Hoffman, supra at 141.

USE OF ALTERNATIVES TO ANIMAL TESTING[*/]

Introduction

The objective of this section is to use an example to determine under what circumstances can data from so-called alternative tests be substituted for those from the traditional whole animal safety evaluation bioassays. Consideration has been given to a number of documents referring to FDA policy-making activities that have, in the past, been the subjects of the Agency's informal rule-making procedures:

a) The threshold assessment of the sensitivity of methods (SOM) policy;
b) FDA's contaminants policy;
c) The so-called red book;
d) The cyclamate decision; and
e) Other FDA decisions.

Examination of such documents revealed that in the past ten years, the FDA completed a number of informal rule-making proceedings and that the issues decided required the analysis and evaluation of experimental data from traditional and from alternative types of tests. The experimental data considered

---

[*/] This Section was prepared by Constantine Zervos and edited by Meyer, Faller and Weisman, P.C.

during these proceedings, some of which began nearly a decade before they were concluded, were collected in the sixties and early seventies. Not surprisingly, therefore, neither the administrative nor the hearing records - whenever hearing records were established - contained extensive references to alternative tests, their evaluations, or indications of the weight, if any, given to the information obtained from them. Simply, the collection of data reviewed in these proceedings antedates the scientific concept of using alternative tests to assess toxicity. Indeed, a cursory review of the cyclamate and aspartame decisions reveals that alternatives to higher mammal testing for cancer were not considered at all. Further, data from what have become known as alternative tests that happen to have been reported in the literature at the time of the proceeding and decision, were given marginal consideration only, to determine whether they support conclusions already reached on the basis of data from the more traditional form of testing. 42 Federal Register, 54148; 41 Federal Register, 21902; 43 Federal Register, 25658.

Thus, it appears that the vast majority of the completed rule-making proceedings of the FDA do not provide adequate examples, suitable for the study and discussion of the circumstances during safety evaluation under which it is permissible and advisable to consider substituting alternatives to higher mammal testing. In one FDA document, however, the Threshold Assessment (TA) of the Sensitivity Of Methods (SOM) regulation, the scientific foundations of safety evaluation were articulated, proposed, and critiqued by the public, and finalized after the concept of alternative tests for cancer became of age. 44 Federal Register, 17070. The TA therefore, is supremely suited to be the example of this section.

I. RISK ASSESSMENTS OR SAFETY EVALUATIONS: SCIENTIFIC SPECIFICATIONS OF STATUTORY REQUIREMENTS

It should not be surprising that the legislative mandates to be enforced provide the blueprints for the scientific frames of reference for public health and environmental protection decisions. They impose upon all involved in the process a number of statutory requirements, which translate into regulatory questions that can be answered only by the scientific method of inquiry.

A specific and well known example of such requirements is the "DES proviso" of the Delaney Amendment to the Food Drug and Cosmetic (FD&C) Act (21 USC 512), which will be further discussed below in connection with the Threshold Assessment. It states that a carcinogenic agent proposed for use as a drug for food animals, can be approved only if, after use is proposed, it leaves "no residue" in the animal tissues of treated food animals. Obviously, whether an agent is a carcinogen and whether its residue is in edible tissues are scientific questions and the intent of the statutory scheme of the Act is to use experimental data to answer them. See, Hess & Clark, Div. of Rhodin, Inc. v. FDA, 495 F.2d 975 (D.C.Cir. 1974); Chemetron Corp. v. U.S., 495 F.2d 995 (D.C.Cir. 1974).

In the absence of specific statutory directive language on how to establish that a substance is a "carcinogen" or that there are "no residues" of it in a particular tissue, those determinations are left to the Agency's discretion. This is the prevalent coursein health and environmental protection statutes. The difficulties inherent inthis concept are exacerbated when statutory requirements are themselves very general, such as that of proving "safety". A regulatory agency considers discernible legislative intent for guidance in applying its internal scientific and technical expertise.

The safety determinations that have evolved from statutory requirements reflect one general underlying principle: given an agent and population exposed to such agent in a specified way, only an empirically-defined potential of agent-caused harm to the collective health of the population is acceptable. As a consequence, the major difficulty with safety requirements arises from the fact that because, it is possible experimentally to observe and measure potential hazardous agents, but impossible to observe their absence.

A serious dilemma arises when a law requires that an agent be demonstrated to be safe before it can be approved for use. There are no scientific tools that can allow one to make observable measurements that can be directly interpreted to mean that a specified agent (e.g., a food additive or a human drug) will not produce harmful effects to those exposed to it. The available experimental tools permit only observation of some of the harmful or beneficial effects of exposure to the agent, and estimation of the probabilities that such effects will follow such exposure. Appropriately interpreted, such observations and estimations lead to assessments of some of the risks of exposure to the agent. However, if, during an experiment, scientists cannot observe an effect that can reasonably be attributed to the agent of concern, the strongest possible prediction they can make about that agent is that it is unlikely to cause the sought after effect under similarly contrived circumstances of exposure. Such a prediction is dramatically weaker than predicting that the agent will have no effect under any circumstances which is the necessary and sufficient condition for safety if the term is given its dictionary definition.

This dilemma with predictions about the potential adverse effects of exposure to agents that must be proven safe before marketing, is resolved through the use of scientific specifications that comply with the statutory scheme of using experimental data to decide whether the requirements of a law have been met. Specifications belong in three categories:

a) Operational definitions of legal standards;
b) Procedures of collecting and reducing data; and
c) Criteria of quality and appropriateness that ensure that applicable operational definitions have been met.

1) Anatomy of Safety Evaluations:

Health and environmental protection statutes contain many key requirements, which, just like the "no residue" requirement of 21 USC 512(H), must be defined operationally to be practically meaningful. They relate to the two central statutory themes of legislation of the genre, "safety" and "efficacy".

The process of constructing a scientific frame of reference for regulatory decisions implementing such requirements, comprises the following steps:

a) Identification of those features of the statute that require scientific specification in order to be implemented;
b) Interpretation of those features in the light of current scientific knowledge and technical capabilities; and
c) Development and articulation of the operational definitions, the procedures, and the criteria necessary for public health or environmental protection, consistent with the governing statute.

II. BASIC SCIENTIFIC SAFETY EVALUATION THEORIES

The science of toxicology operates under three scientific theories about the possible hazardous or toxic effects of regulated products: (1) The receptor theory of toxicity; (2) The chemical reactivity theory of toxicity; and (3) An older theory of toxicity which holds that in general, biological systems are less sensitive to the effects of chemicals than physical systems; therefore, if it is impossible to detect an insult or a potentially insulting substance, the substance cannot be toxic at undetectable levels of hazardous concentrations and therefore, should not be considered to be hazardous.

For many important reasons, scientific theories cannot be applied directly as regulatory requirements in the construction of operational definitions. In addition to being credible, standards for safety evaluations must also be sound, i.e., valid over protracted periods of time. Operational definitions based on scientific theories would have to change to reflect changes in scientific theory. Frequent changes in regulatory requirements create uncertainty and chaos in the marketplace. The uncertainty as to which theory is best is generally considered detrimental even from the public health protection view. Regulatory activities based on safety evaluations of transient validity are usually perceived by the public as unreasonable government hindrances to technical development. The use of animal tests as the principal basis for assessing the potential risks from exposure to toxic chemicals has, however, been vigorously upheld by the courts. EDF v. EPA, 510 F.2d 1292 (D.C.Cir. 1975).

III. FDA OPERATIONAL DEFINITIONS OF SAFETY

The FDA uses a number of operational definitions to uphold the safety standards in the statutes it enforces. Generally, they fall into one of three categories which relate to the three scientific concerns about toxicity.

Which of the three categories applies to a specific substance depends on the intended use of the substance (i.e., drug, food additive, etc.) and on the toxic effects the substance is capable of eliciting. Traditionally, if a substance causes organ toxicity, e.g., some very special forms of cancer only, or fetal toxicity, it is regulated according to operational definitions of safety that derive from the combination of the receptor theory of toxicity and from either one of two axioms: (1) that no risk is acceptable, or; (2) the acceptable risk must be commensurate with the expected benefits from the proposed use of the substance.

Where a regulated product must present no risk, strict operational definitions of safety are applied to the approval process for food additives and non-carcinogenic animal drugs. In the latter category belong regulated products, e.g., human drugs, biological products, devices and food contaminants. If a substance causes cancer, mutations or degenerative effects, it is regulated according to operational definitions of safety that derive from the chemical reactivity theory of toxicity, and again from either one of the same two axioms. However, the operational definitions of safety that exclude carcinogenic substances from food additive uses, permit the uses of carcinogenic substances as food animal drugs.

Below is a sample of operational definitions of safety currently used by the Food and Drug Administration.

a) Operational definition of safety for food or color additives which are not carcinogens:

A potentially toxic substance which is not a carcinogen can be used safely as a food or color additive if the anticipated maximum human exposure to it, including the exposure from the proposed use, is less than a fraction of the exposure found to cause no observable adverse health effects (NOEL) in experimental animals under specified conditions of bioassay;

b) Operational definition of safety for animal drugs which are not carcinogenic:

A potentially toxic substance, which is not a carcinogen, can be used safely as a food animal drug if the maximum anticipated human exposure to it in the form of residues in the edible tissues of treated animals is less than .1 part per million;

c) Operational definitions of safety for food and color additives which are carcinogenic:

A substance which is carcinogenic cannot be used safely as a food or color additive under any circumstance. This is the famous or infamous Delaney clause of the Food Drug and Cosmetics Act.

d) Operational definitions of safety for animal drugs which are carcinogenic:

A substance which is a carcinogen can be used safely as a food animal drug only if it leaves no residues in the edible tissues of treated animals.

There are a number of other operational definitions of safety used by the Food and Drug Administration. They are more complex and they apply to drugs, medical devices, pesticides, biological products, and vaccines. In general, these operational definitions of safety share a common feature that the accepted risk of drug-induced adverse health effects is somehow judged to be preferable to not having the benefits of the uses of the substance or the product.

IV. CHANGES IN THE REGULATORY SAFETY TESTING REQUIREMENTS

A change or shift from the traditional whole animal tests to available alternatives will require a change in the relation chosen to predict toxic effects as a function of exposure to a substance under study. This is equivalent to changing the applicable operational definition. But, scientific laws and axioms, the other two pillars that support operational definitions of safety are not easily changeable at the discretion of scientists, administrators, or legislators, shifting from one type of test to another is equivalent to changing the basic scientific assumptions.

Federal agencies do, from time to time, change their policies. One would therefore like to review several changes of the genre to see how they come about and what factors contribute to success. However, as a matter of rule, appearances of discontinuity in policy are avoided. Thus, usually even the most abrupt policy changes are presented as matters of policy evolution, and official descriptions of the factors and circumstances that contribute to change are not always clear or reliable.

In the recent past, the FDA has attempted to institute two policy changes that bear on the objectives of this document. They are the changes in the operational definition of the "no residue" requirement of 21 USC 512(H), and the Threshold Assessment (abbreviated as TA). Fortunately, the Agency made a sustained effort to present these activities to the public exactly for what they are: abrupt changes in policy. In both cases therefore, there is a record that can be studied.

Although for all practical intents and purposes the new definition of "no residue" has been in effect since the middle 1970's, the final rule that institutes the policy change has not yet made it through the informal rulemaking process of the Agency. This is one reason why it was not selected as the example of this study. The other is that the change exemplifies a trend toward greater dependence of whole animal tests no less. Under any circumstances, this change has been analyzed in some detail. (Zervos, C.;"Food Safety; Interface Between Legal Requirements and Analytical Possibilities" in Instrumental Analysis of Foods, Vol. II, 1983, pg.1; Charalumbous and Inglet, Editors). From this analysis, which is the prototype for the T.A. treatment below, it is easy to discern the forces that increase our dependence on whole animal tests because of the nature of the alternative definition of the "no residue" requirement. However, it is not easy to see the forces that decrease that dependence. The latter forces become more apparent in the case of the T.A.

A. The Threshold Assessment: Background:

A choice between two alternative approaches to an operational definition of safety, one requiring extensive use of whole animal tests, and the other based on "alternative" tests, was made by the Food and Drug Administration in the case of animal drugs.

In the discussion of the various types of operational definitions used by the FDA, it was noted that carcinogenic animal drugs are regulated according to the chemical reactivity theory of toxicity, and the other, according to the receptor theory of toxicity. It was also noted that this policy of dual control is based on the Agency being able to tell substances which pose a carcinogenic risk from those that do not. Early in the evolution of its policy in implementing 21 USC 512(H), the FDA faced the question of whether to require the protracted and expensive whole animal tests for carcinogenesis for all substances proposed for use as animal drugs. Alternative ways were sought to attain the same degree of safety and public health protection without the expense and time requirements of carcinogenesis bioassay test. See, Food and Drug Advisory Committee on Protocols for Safety Evaluation of Food Additives and Pesticides, Toxicology and Applied Pharmacology, 20:414 (1971).

B. The Choices:

The Threshold Assessment is a statement of FDA policy concerning the requirements for testing potentially carcinogenic chemicals proposed for use in food-producing animals. Specifically, it spells out the circumstances and the conditions under which data from metabolic and pharmacologic studies, as well as data from structure-activity correlations and from short term toxicity tests can be substituted for data from cancer bioassays. It is the product of FDA's informal rulemaking procedure (21 CFR 16). It was proposed for public comment several times (38 FR 19226; 42 FR 10412; and 44 FR 17070) and each time it was modified to accommodate a number of public comments received in response to an agency request for such comments.

The FDA decided all outstanding issues related to the Threshold Assessment without a hearing before an administrative law judge. Thus, there is no hearing record. However, there is an administrative record (reference) which contains the transcript of a two day public hearing before FDA officials held for information-gathering purposes on June 22 and 23, 1979. From the administrative record, the many comments submitted by members of the public, and from FDA's responses to them (i.e., from the informal rule-making process), it is possible to identify and analyze the scientific basis and other relevant reasons for deciding under what circumstances available alternatives to higher mammal testing are permitted and adequate. From the same record, it is also possible to draw conclusions about the implications of accepting or rejecting alternative tests.

The FDA began to articulate its food animal drug policy in 1973. It then proposed to consider data from tests other than those obtained from whole animal bioassays in order to lighten the burden of testing. Thus, in the first proposal of the agency's SOM policy, the agency proposed to exempt from whole animal carcinogenic bioassays which, in the opinion of FDA scientists and the commissioner, could not possibly present a significant risk of carcinogenesis to the consuming public

Although initially the questions of significant risk was left vague, the FDA through three cycles of proposed public comment, became more specific. The agency devised an alternate operational definition of safety for animal drugs that are known neither to be carcinogens nor non-carcinogens. This alternative operational definition is founded on data from three other sources of data: (1) structure activity correlations; (2) metabolic data that bear on the concentration of animal drugs on edible tissues of treated animals, and; (3) data from so-called short term *in vitro* tests of genetic toxicity. Ultimately, the agency devised a protocol for assigning numerical scores to substances of unknown carcinogenic potential proposed for use as food animal drugs. It is essentially a linear combination of particular scores assigned to a substance according to the available amount and the nature of data from the three specific sources. If the score of a substance is found to exceed a specific value, then the substance must be subjected to the rigors of whole animal carcinogenesis potential assessment and tests reserved for substances that are carcinogens or suspected of being carcinogens. If, on the other hand, the numerical score of the substance is less than the discriminating value, then the substance is considered safe for food or animal drug use if its level of residues in edible tissues from treated food animals is less than a specified concentration, e.g., one-tenth of one part per million.

The FDA, in rejecting a generalized requirement for whole animal testing and accepting an alternate operational definition of safety under certain circumstances, the agency went through a detailed and protracted analysis of all factors that could be imagined by agency staff and all those suggested by many people who commented on the various FDA proposals. The agency went through long and agonizing analysis and discussion of the relative levels of public health protection conferred by the two alternate operational definitions of safety, namely that which requires whole animal cancer bioassay tests and that which does not. The agency considered the equity of the two procedures and found them satisfactory. It considered the administrative efficiency of the procedures and found the alternate operational definition to be less cumbersome and less demanding.

The FDA considered at length the legality of the two alternative operational definitions, and concluded that both were legal under its statutory authority. The agency considered, of course, whether either one of the two operational definitions faced any unusual cultural barriers, it considered the scientific credibility of the two alternatives, the technical complexity, and also, finally, whether adopting either one of the two alternatives would have untold side effects in terms of the agency's ability to carry out its legally mandated tasks.

After consideration of the issue and the two alternatives for the better part of a decade, the agency concluded in 1981, that the operational definition and the basic underlying regulatory evaluation did not require whole animal tests and in vitro alternatives are preferable, it finalized as guidelines the Threshold Assessment (TA) policy. See FDA Guidelines and Procedures for Evaluating Assays for Carcinogenic Residues, December, 1981. It should be noted that this change of policy on the part of the Food and Drug Administration had absolutely nothing to do with the rise in prominence of humane cancers. It came about as a result of the realization that whole animal experiments are expensive and the fact that after thorough analysis, they do not seem in certain situations to provide any greater degrees of assurance about the absence of toxic effects than other tests currently available.

ALTERNATIVES TO ANIMAL USE IN TESTING AND EXPERIMENTATION

Chapter Nine

Economic and Policy Considerations

Final Report

Prepared for the

Office of Technology Assessment
U.S. Congress

Prepared by

Henry R. Hertzfeld
Washington, D.C.

and

Thomas D. Myers
McLean, Virginia

January 18, 1985

TABLE OF CONTENTS

ALTERNATIVES TO ANIMAL USE IN TESTING AND EXPERIMENTATION: ECONOMIC ISSUES

Page

SECTION I: COSTS

INTRODUCTION 1
APPROACH 1
PREVIEW OF THE CONCLUSIONS 2
COSTS OF SHORT-TERM IN VITRO TESTS 4
COSTS OF A REPRESENTATIVE SAMPLE BATTERY OF IN VITRO TESTS 6
COSTS FOR A BATTERY OF CHRONIC WHOLE ANIMAL TESTS 7
SHORT-TERM VERSUS CHRONIC LONG-TERM TESTS 9

SECTION II: RELIABILITY

VALIDATION AND CORRELATION 19
ASSAY PERFORMANCE OF THE AMES SALMONELLA REVERSE-MUTATION ASSAY 21
PERFORMANCE OF THE AMES SALMONELLA ASSAY IN BATTERY 23
CORRELATION COMPARISON OF THE RODENT AND AMES BACTERIAL TESTS 24
AMES SALMONELLA ASSAYS - LIMITING FACTORS 29
  Chemical Classes 29
  Potency Data 29
SUMMARY 30

SECTION III: IMPACT OF TECHNOLOGICAL CHANGE

ECONOMIC CONSEQUENCES 31
  Overview of the Industry 31
  Relationship of Research and Development to Testing 32
  Toxicological Testing Services 33
  Implications for the U.S. Budget 37
POLICY CONSEQUENCES 40
  Implications to Risk Assessment Methodology 40
    Expanded Screening Applications: 41
    Substitution for Animal Tests: 42
  Statutory Limitations and Regulatory Constraints 43
  Public Perception 45
  International Implications 46
  Summary 46

TABLE OF CONTENTS
(Continued)

Page

SECTION IV: ECONOMIC ANALYSIS OF BASIC AND APPLIED RESEARCH

INTRODUCTION 48
THE PROCESS OF TECHNOLOGICAL DEVELOPMENT 50
ECONOMIC MODELS 51
  Uncertainty, Risk and Benefit 51
  The Measurement of Economic Risks and Benefits in Emerging Technologies 53
  Limitations and Philosophical Concerns 56
ECONOMIC MODELS OF BASIC CONCERN 57
CASE STUDY: ALTERNATIVES TO WHOLE ANIMAL TESTING 60
  Two Economic Risk Models Applies to Alternatives to Whole Animal Testing 63
CONCLUSIONS AND RECOMMENDATIONS 68

TABLES

Table II-1 22
Table II-2 26
Table II-3 27
Table II-4 28

Table III-1 34
Table III-2 35

Table IV-1 64

ALTERNATIVES TO ANIMAL USE IN TESTING AND EXPERIMENTATION: ECONOMIC ISSUES

SECTION I: COSTS

INTRODUCTION

Over the past fifteen years, various short-term in vitro tests for mutagenicity and carcinogenicity have been developed that have the potential for replacing the traditional tests that use whole animals. These in vitro tests are relatively inexpensive and yield results after a very short period of time.

The most popular, and one of the first such tests introduced, is the Salmonella typhimurium/microsome plate mutation assay (the "Ames test") (10). This assay, which has been in use less than fifteen years, has the most extensive data base developed (4). Although it may not be the cheapest nor the best test for mutagenicity, it is the focal point of this economic analysis primarily because the data exist and an analysis is possible.

It should be emphasized that the analysis will not be limited to the Ames test alone. Most recent work has recommended that a battery of short-term in vitro tests be used (23). This battery usually starts with the Ames test and includes four or five additional tests (36). The following analysis will focus on the various recommended series of tests.

APPROACH

Problems inherent in addressing new technologies, as well as the economic measurement problems encountered in any data series, preclude precise measurement of all economic factors. There are several reasons measurement is particularly difficult when dealing with new technologies. One is that data are often proprietary and held closely by the producing companies. Another is that the Federal data sources have not caught up to the technology and have not yet adequately redefined industries to separate out the new from the traditional processes and products. Also with rapidly changing production processes, it is likely that no two firms use exactly the same processes, and comparisons are not at all exact.

In addition to the above problems of new technologies, there are economic measurement problems that exist with any data series. These will not be elaborated here, but they include problems with survey sample and responses, inadequate detail for measuring smaller industries, disclosure and confidentiality problems, improper classification of firms producing many different products, etc.

We have attacked the economic cost estimating for the comparison of short-term in vitro assays with chronic whole animal testing procedures in the following manner. First, there are several existing summary studies that include cost and price data for most of these assays and which are relatively recent (post 1980). These are the basic reference sources. Second, there are current price lists for the various laboratories in the toxicity testing industry. These have been reviewed. Third, we have spoken with representatives in the industry to check the data in the various survey reports. From these sources, an economic profile of the industry has emerged.

PREVIEW OF CONCLUSIONS

The economic case for using short-term in vitro bioassays instead of chronic whole animal testing is compelling. Testing, whether it be in vitro or in vivo, is primarily human effort and labor (estimated to be 80% of the costs) (11). Therefore, differences in materials and equipment between types of tests and procedures becomes relatively unimportant to the analysis compared to the elapsed time and the expensive human labor expended over long testing protocols. A test, or a battery of tests, that take one to three months and cost somewhere between $25,000 and $50,000 to perform will be far more economic and cost effective than a battery of tests that take three to five years and cost upwards of $1,000,000.

The crucial issues will not be decided by a simple economic analysis of assay cost comparisons. The important factors are reliability of the tests and public acceptance. In other words, the scientific merit of the tests and the political process. We believe that over a long time span (ten to twenty years), as the technology develops and matures and a sufficient data base is acquired to adequately measure the variables, we can now only assess with vague probabilities, that economic factors will be decisive in directing the technological development and market acceptance. These factors will be much broader than the relative cost comparisons of testing procedures. They will encompass all aspects of societal costs and benefits from the use of newer technologies involved in toxicological testing. But over the next few years, economic arguments will take second place to the scientific and regulatory community's arguments for and against using these short-term tests as replacements for whole animal tests.

The most important current economic issue is collecting appropriate data, analyzing it correctly and structuring a quantitative approach. At present, with the technology in its formative stages, quantitative analysis will require substantial judgment; therefore, qualitative arguments will have to be substituted. As the data base becomes much more extensive in the future, it should be possible to substitute economic arguments that will have greater policy relevance.

COSTS OF SHORT-TERM IN VITRO TESTS

Comparative cost analyses are difficult. In the United States, there are over 110 organizations performing toxicological testing, including universities, in-house industrial laboratories, and commercial testing facilities. The mix includes not-for-profit as well as profit making institutions. Accounting procedures for capital expenditures, overhead and labor are very different among the various organizations.

Even among similar institutions, such as commercial testing laboratories, the procedures, protocols, and other factors that are inputs to the testing process for regulatory and other purposes vary considerably. Take, for instance, the cost differences to a laboratory of having in-house employees read test results on slides versus a laboratory sending the slides out to individual subcontractors who read the results in their homes. The former method requires hiring employees, paying fringe benefits, and other overhead and fixed costs, while the latter involves only paying for the services as they are needed.

Another variability factor is in the range of tests that are offered. Some labs specialize in one or a few tests and do not offer a full range. Some are oriented heavily or solely toward animal testing, while others offer both animal and molecular testing. Cost accounting systems will allocate particular items of overhead and depreciation very differently.

Any particular test cannot accurately be compared across different labs for cost of production. Not only do the above variables differ among different labs, but each lab may recommend a slightly different version of the same test, or a slightly different battery of tests. Protocols, number of dose levels, etc. may vary considerably. In doing an Ames test, for example, variables include the number of strains, the number of dose levels, whether a S9 activation is used or not, and whether it is done with or without preincubation (25). Detailed survey data are not available from lab to lab to standardize the cost comparisons. And even if one could standardize the costs, it would be difficult to compare labs because the scientific community is not in agreement as to which of the variables is optimum for a given purpose (i.e., testing a particular chemical for mutagenicity and carcinogenicity).

Some reasonably accurate cost comparisons can be made. Price charged a customer is a good proxy for cost. The toxicological testing industry is currently competitive, with four or five large labs in the U.S. offering the test to industry on a profit basis. In addition, there are over 80 other firms also doing molecular and/or animal testing (47). During the late 1970's many firms were gearing up for increased testing to meet regulatory requirements. In the early 1980's, the anticipated demand did not materialize and the industry is now operating at 60% to 70% of capacity. Therefore, competition has resulted in prices that reflect the costs of production (47).

A price quotation for the Ames test or for any other test may not reflect costs of production, as it can be offered as a loss leader, or it can be costed jointly with other tests. A more accurate measure would be the comparison of a series or battery of tests. This should be noted carefully when extracting any partial results from the tables included with this report.

Common to the entire cost/price analysis is the problem of the quality of the work done by the testing lab. The economic data available for the various firms makes no reference to the firm's reputation for reliability, quality, timeliness or reports, etc. Over a long run, a firm that is unreliable will lose business and drop out of a competitive market (it will cost customers more to redo results, effectively raising that firm's price). There is no accurate way of quantifying quality differences although some insight into this might be available through the results of laboratory audits conducted under the Good Laboratory Practice Regulations.

COSTS OF A REPRESENTATIVE SAMPLE BATTERY OF IN VITRO TESTS

The following table represents the battery of tests now generally recommended by the National Toxicological Program for testing mutagenicity and carcinogenicity (36). The first five tests are generally recommend; the sixth is often optional.

	Test	Cost/Price Range			Mid-Point
1.	Ames Test	$ 900	to	$ 2,335	$ 1,618
2.	Mouse Lymphoma	4,500	to	11,225	7,863
3.	SCE-CHO Cells	2,500	to	6,000	4,250
4.	Drosophilia	7,000	to	20,850	13,925
5.	Cell Transformation BALB 3T3	4,400	to	5,500	4,950
6.	Unscheduled DNA Synthesis	4,700	to	5,500	5,100
	Total	$24,000	to	$51,410	$37,705
	Total without #6	$19,300	to	$45,910	$32,605

A number of factors must be emphasized. The range of prices includes: actual prices charged during 1984 at different labs, protocol estimates made in 1980 for the EPA by a contractor, 1980 survey results, 1983 survey results (12,17,16,18,23,48,50,55). No attempt to standardized the figures was made, nor were the data adjusted for inflation. The reason these adjustments were not attempted were as follows: the range was very broad, with some 1984 prices below 1980 estimates for particular tests; protocols were differently interpreted for many of the individual tests, making exact comparisons impossible; and, as detailed above, there were myriad dissimilarities among the structure of the organizations doing the studies. Fine tuning non-standardized data would be fruitless, since all the data for specific assays fall within a broad but comparable range.

The Ames test price ranges as listed above include prices for both conducting the test once and for an independent repeat of the test. The low end (around $1,000) for the Ames test represents the one-time test, which for most EPA and FDA reviews is adequate. The high end (around $2,000) is for the Ames test plus the independent repeat, which may be needed in some circumstances.

We have been told that an upper limit for prices for a battery of five tests is $25,000, and that an average customer pays about $10,000 to $12,000 (11). This indicates that there may be significant "package" deals for a battery of tests done in one laboratory, and that simply adding up the figures as we have done may be overstating the costs and price of a representative battery. Further, if the average customer pays less, it is quite likely that, for any particular chemical, the recommended battery may not include all the above tests, but only a few, with more testing to follow if the initial tests indicate the need for further information.

COSTS FOR A BATTERY OF CHRONIC WHOLE ANIMAL TESTS

Chronic testing of whole animals for mutagenicity and carcinogenicity is a lengthy and expensive process. To meet regulatory requirements, it may be necessary not only to test one species, such as rats, but correlative tests may be required for mice, hamsters, and perhaps dogs as well (12). Usually both sexes must be tested and several generations observed. This procedure can take from two to five years (12).

The Enviro Control, Inc. study for the EPA in 1980 analyzed in great detail all the elements that make up a testing procedure (12). Their protocol estimates, which focus on in vitro testing, take into account labor, overhead, cost of lab materials, general and administrative costs, and the profit. Each category is broken into its component costs, and costs as detailed as the rations, bedding sheets, watering, and cleaning of the cages for the animals are estimated. These investigators predictably ran into difficulty in estimating the components of the overhead figure. Using an average of 115% of direct labor, they avoided the difficult issue of allocating fringe benefits, depreciation and/or rental fees for the laboratory, etc. They observed that individual firms used vastly different methods of calculating overhead and that they were not capable of striking a representative estimate (12).

Analyzing their data, we find that physical materials (including the procurement of the animals) used in the lab average just over 10% of the total cost of the testing procedure. For chronic combined rat, mice, and dog tests supplies ran 12% of the $1,000,000 protocol estimate. For a much simpler acute oral toxicity study, materials were 10.6% of the $2,400,000 protocol estimates.

It is assumed that fringe benefits are 30% of the overhead, and this is added to labor costs, then we have estimated that somewhere between 50% and 90% of the total cost of a test is related to human labor, including administrative costs and management fees. In conversation with experts in testing laboratories, an 80% figure for labor related expenses was quoted, which appears to be a reasonable figure, given the data in the Enviro Control report (12).

These figures are important because it emphasizes the importance of time to the testing procedures. The more the testing is dependent on human labor (and both animal and in vitro testing are labor intensive), the greater the savings that will occur with reduction in the testing time frame.

A set of tests that requires several years will cost a great deal more that a set that can be finished in several months. Estimates of the costs of chronic whole animal tests can vary as much as those for in vitro tests. The minimum figure quoted was $325,000 (in 1980--since then inflation will have raised this to at least $400,000). High figures, again for 1980, range to $1,600,000. Today, estimates range between $500,000 and well over $1,000,000 for these tests (12). Specifics are not well defined in the published price lists from the various laboratories. They do not quote prices for these chronic whole animal tests. Instead, the prices appear to be negotiated in each case (17,16,18,48,50).

Laboratories that do whole animal testing that can take up to several years for each chemical tested run additional economic risks. Prices are customarily quoted before the testing begins when a contract is made with the client. If inflation drives up salaries and materials faster than they have projected, the laboratory could lose money. The trend in prices for chronic whole animal tests has been upward. This could, in part, be accounted for by the period of rapid inflation in the late 1970's and early 1980's which may have adversely affected the profit position of many of these labs.

Conversely, the laboratories that do short-term in vitro testing are not without risk. If an average test takes about 90 days, then they are dependent on a steady stream of fast turn-around work. Without it, they will have to lay off personnel and make adjustments. Fixed costs, such as buildings and equipment would be a financial burden until business improved. An individual firm specializing in short-term tests in a very competitive and changing market could find a significantly higher degree of business risk that a firm that does all types of testing.

SHORT-TERM VERSUS CHRONIC LONG-TERM TESTS

Each type of testing, both short-term and long-term, has its advantages and disadvantages from a corporate planning, pricing, and profit perspective. In the abstract, it is not to say which is a more profitable type of business, as macroeconomic conditions coupled with regulatory policy decisions will dictate the important parameters (inflation, number of chemicals backlogged, acceptability of the various testing procedures, etc.).

The cost estimates reviewed above are per chemical tested, as are the costs reviewed for the battery of short-term tests. From an economic efficiency view, assuming equal reliability from both series of tests, the short-term tests have an advantage. Not only can industry more quickly and more cheaply get new chemicals on the market, but from a societal perspective, this permits many additional drugs, chemicals, and other agents to be tested with the same resources or those resources not consumed by expensive tests to be put to use elsewhere in the economy. As the next section will elaborate, the assumption that these short-term tests are equally reliable is not presently accepted by the majority of the scientific and regulatory community; therefore, policy analysis based on economic efficiency arguments is premature.

The economic impact of the newer in vitro technologies in testing is illustrated by the following examples. Over 50,000 chemicals are now in use that have not been subjected to testing procedures (36). This backlog is aggravated by the introduction of about 500 new chemicals per year. (The actual rate of introduction depends on many factors including the economy, market projections for specific substances, the technological advances in the chemical, drug, and pesticide industries, as well as the current regulatory "climate".)

Currently, somewhere between 100 and 200 chemicals get tested each year. If the toxicological testing industry was working at full current capacity, the maximum number of chemicals that could be subjected to tests would only be on the order of 350 per year (7,47). Clearly, many chemicals will not be tested using the current testing protocols. Priorities have to be assigned based on estimated usage, estimated potency, and various screening assays and tests. From a social perspective, this can and will be costly, as mistakes and false results inevitably will occur. Given that no testing system now known is infallable, society's major concern should be how to test as many substances as possible and minimize errors. The constraint is the amount of resources society is willing to devote to this testing.

Let us assume that the current funding level represents society's choice for investment in toxicological testing and that this "buys" tests for about 200 substances per year, with most of the testing consisting of chronic whole animal tests. The cost for this, if an average chronic testing protocol is about $1,000,000, is $200 million per year. (Note that this is close to the estimated total sales for the commercial labs in the U.S. per year. It is likely that the $650 million total toxicological testing industry figure mentioned in other parts of this report represents a sizable amount of research as well as testing. Corporate in-house laboratory's allocation to testing versus research are difficult to separate as both will go on under the same roof.)

We further assume that an average battery of short-term in vitro tests will cost $30,000 per chemical. Just under 7,000 chemicals per year could be tested with resources of $200 million per year allocated totally to in vitro testing. The backlog of untested substances could be eliminated within eight years, with society not spending any additional resources than are now spent annually.

If these in vitro tests were to be proven as accurate as the current chronic whole animal tests, and if regulatory agencies accepted the tests as conclusive (both highly unlikely occurrences in the near future), then it is probable that the increased demands for these in vitro tests would drive the industry to develop even better testing technologies at a faster rate. It is reasonable to speculate that not only better batterys will be discovered but that the tests will become increasingly standardized, thus permitting additional use of computer techniques for monitoring and performing the tests. These automated processes should substitute for expensive human labor. With high demand for the testing, there will also be scale economies developing and all the factors acting together will drive the price of such tests down.

Looking to the future and assuming that the above scenario is feasible, it would not be unrealistic to assume a price of $10,000 per battery of tests per substance, a time frame of performing the tests of a month or less, and a clearing of the backlog of untested substances in only a few years. And, a stimulation of the chemical, drug, pesticide, and other affected industries would occur, permitting many companies to experiment and bring to market new and safe substances at significantly less cost than is possible at present.

The above hypotheticals represent extremes. Today's practices are expensive and incur significant private and social costs. They also produce benefits in that some dangerous substances have been controlled. Tomorrow's technologies offer the promise of being much cheaper and increasing benefits without incurring additional costs, and, in extreme cases, saving significant resources.

These examples, in order to illustrate a point, have overlooked many factors. On the economic front, if all testing were to shift to in vitro assays, then many firms now geared to animal testing would find their labs practically useless. New testing equipment and technologies could even induce new firms to enter the market and shift business from current labs providing in vitro testing to new companies with rights to as yet undeveloped and superior testing methods.

The status quo is a strong motivator for those now profitable firms. Regulatory authorities tend to follow, not lead, in accepting new technologies, and they are heavily influenced by the industrial concerns now in place. Economic realities eventually prevail, but the short-term turmoil in the industry and regulatory agencies will be painful to many. The danger is that a new technology that could lead to significant social benefits will be adapted with a long delay (and increase social costs by a large degree) because of political and institutional barriers. Congress and the government should be alert to these possibilities.

Social costs and benefits are difficult to evaluate. They are based on myriad assumptions, expert opinions, statistical avarages, and other error prone judgments. Lave, Omen, et al. have argued the economic costs to society of inaccurate (either false negative or false positive) results can be very large (17). Their assumption, that the average social costs of incorrectly classifying a carcinogen are $10,000,000 and of incorrectly classifying a noncarcinogen are $1,000,000, is unexplained and open to debate. False positives not only prevent industry from making profits and bringing new products to the market, but they may also incur significant societal costs. For instance, withholding a new drug from the market place because tests considered it carcinogenic when it really was not could result in considerable costs to the population that might benefit from having the drug. It all depends on the relative number of people affected. Conversely, if a drug was considered noncarcinogenic from the tests and it really was a carcinogen, the social costs could be low if the number of people who might have to take the drug is relatively small.

The authors strike averages. In doing so, they obscure the particulars of specific cases. And they do not adequately explain their assumptions. A $10,000,000 social cost to a false negative finding appears low. If we take a hypothetical use where permitting a chemical on the market results in significant harm to a reasonably large population, it is extremely hard to calculate the value of a human life or the monetary damages that an injured person suffers. In spite of the difficulty, courts and insurance companies make such calculations every day. Awards of hundreds of thousands and several million dollars per person are common, particularly if the person is young and/or of a high income potential. Thus, the author's "high" average figure of a $10,000,000 societal loss may be quite low. Nonetheless, the numbers they calculate are instructive and illustrate the vast potential societal costs of inaccurate testing. Section IV will pursue the use of such economic models in more detail.

File: OTA.COST.EST
Report: BIOASSAY COMPARISON

Test	Cost/Price	Range	Data Type	Resp	Source	Date
Reverse mutation in E. coli Strain WP2	300	n.a.	Actual	1	Litton Bio.	1984
Escherichia coli, WP2	400	n.a.	Survey	2	NAS, Tab. 17	1982
Skin Irritation, rabbit	425	200 to 700	Survey	8	van der Lee	1983
Eye Irritation, rabbit	531	200 to 1,050	Survey	8	van der Lee	1983
Acute toxicity, Daphnia magna	625	600 to 650	Survey	2	van der Lee	1983
Bacterial DNA Repair (E. coli, B. subtilas)	650	250 to 1,200	Survey	5	EPA Study	1980
Primary Eye Irritiation	650	200 to 2,150	Survey	22	EPA Study	1980
Acute toxicity, Foecilia reticula	725	650 to 800	Survey	2	van der Lee	1983
Growth inhibition, unicellular algae	750	750 to 750	Survey	2	van der Lee	1983
Bacillus subtilis rec-	800	n.a.	Survey	1	NAS, Tab. 17	1982
Biotic degradation, closed bottle test	800	n.a.	Survey	1	van der Lee	1983
Ames Salmonella/Microsome Plate Assay	900	n.a.	Actual	1	Litton Bio.	1984
Primary Eye Irritiation	900	450 to 1,350	Estimate	n.a.	EPA Study	1980
Ames Salmonella	1000	n.a.	Actual	1	SRI	1984
Bacterial Mutagenicity (Ames)	1000	300 to 1,500	Survey	5	EPA Study	1980
Eukaryotic Mutagenicity (Yeast)	1000	700 to 1,200	Survey	4	EPA Study	1980
Mutagenicity, salmonella test (Ames)	1120	850 to 1,300	Survey	4	van der Lee	1983
Ames Salmonella	1200	n.a.	Actual	1	Hazelton	1984
Salmonella (Ames)	1290	n.a.	Survey	10	NAS, Tab. 17	1982
Ames Salmonella w/ Indep. Repeat	1300	n.a.	Actual	1	Hazelton	1984
Reverse Mutation Induction in Yeast Strains	1300	n.a.	Actual	1	Litton Bio.	1984
Ames Salmonella	1325	n.a.	Actual	1	Micro	1984
Ames Salmonella	1350	n.a.	Actual	1	Bioassay	1983
Quantitative Suspension Mutagenicity, Bacteria	1350	n.a.	Actual	1	Litton Bio.	1984
Acute Oral Toxicity	1400	300 to 2,500	Survey	6	EPA Study	1980
Transformation--CHO cells	1400	n.a.	Survey	1	NAS, Tab. 17	1982
Yeast Cell: Saccharomyces Cerevisiae	1400	n.a.	Survey	1	NAS, Tab. 17	1982
Acute Toxicity, oral (rat)	1465	$900 to $2,2	Survey	8	van der Lee	1983
Ames Salmonella	1500	n.a.	Actual	1	Pharmacon	1984
Escherichia coli rec-	1500	n.a.	Survey	1	NAS, Tab. 17	1982
14 day reproduction study, Daphnia magna	1600	n.a.	Survey	1	van der Lee	1983
Ames Salmonella w/ Indep. Repeat	1750	n.a.	Actual	1	Litton Bio.	1984
Acute Toxicity, dermal (rat)	2077	$1,100 to $4	Survey	8	van der Lee	1983
Ames Salmonella w/ Indep. Repeat	2225	n.a.	Actual	1	Micro	1984
Ames Salmonella w/ Indep. Repeat	2335	n.a.	Actual	1	Bioassay	1983
Acute Oral Toxicity	2450	1,200 to 3,6	Estimate	n.a.	EPA Study	1980
Dermal Sensitization	2500	400 to 10,00	Survey	12	EPA Study	1980
Mammalian Cells (Sister Chramatid Ex.)	2500	1,000 to 20,	Survey	5	EPA Study	1980
Somatic Cells (Unscheduled DNA Synthesis)	2560	1,000 to 20,	Survey	5	EPA Study	1980
Degranulation	2500	n.a.	Survey	1	NAS, Tab. 17	1982
Acute Dermal Toxicity	2650	1,400 to 4,6	Survey	6	EPA Study	1980
Sister-chromatid exchange--CHO cells	3000	n.a.	Survey	1	NAS, Tab. 17	1982
Sister-chromatid exchange--mouse	3000	n.a.	Survey	1	NAS, Tab. 17	1982
SCE--CHO	3100	n.a.	Actual	1	Hazelton	1984
SCE--CHO	3200	n.a.	Actual	1	Bioassay	1984
Skin sensitization (gpmt)	3388	1,700 to 7,0	Survey	8	van der Lee	1983
Micronucleus-mouse	3400	n.a.	Survey	2	NAS, Tab. 17	1982
SCE--CHO	3500	n.a.	Actual	1	Litton Bio.	1984
Chromosome Aberrations in CHO	3600	n.a.	Actual	1	Bioassay	1984
Chromosome Aberration in CHO	3650	n.a.	Actual	1	Hazelton	1984

File: OTA.COST.EST
Report: BIOASSAY COMPARISON

Test	Cost/Price	Range	Data Type	Resp	Source	Date
Dermal Sensitization	3900	1,950 to 5,8	Estimate	n.a.	EPA Study	1980
SCE-CHO	4000	n.a.	Actual	1	Micro	1984
Mutagenicity, micronucleus test	4000	2,000 to 5,6	Survey	5	van der Lee	1983
Acute Dermal Toxicity	4250	2,100 to 6,4	Estimate	n.a.	EPA Study	1980
Transformation, C3H10T1/2	4400	n.a.	Actual	1	Micro	1984
Transformation C3H10T1/2	4500	n.a.	Actual	1	Bioassay	1984
Somatic Cell Mutagenicity (mouse lymphonia)	4500	2,000 to 7,0	Survey	6	EPA Study	1980
Micronucleus Rodent Aberrations	4600	$2,400 to $6	Actual	1	Hazelton	1984
Transformation, BALB-C/3T3	4625	$4,400 to $4	Actual	1	Micro	1984
Unscheduled DNA Synthesis, Rat	4700	n.a.	Actual	1	Hazelton	1984
Mouse Lymphoma Mutation Assay	4700	n.a.	Actual	1	Litton Bio.	1984
Unscheduled DNA Synthesis, Rat	4800	n.a.	Actual	1	Litton Bio.	1984
CHO-HGPRT	4900	$4,900 and u	Actual	1	Hazelton	1984
TK +/- L5178Y mouse lymphoblasts	4900	n.a.	Survey	1	NAS, Tab. 17	1982
Unscheduled DNA Synthesis, Rat	4950	n.a.	Actual	1	Pharmacon	1984
Mutation Assay in CHO Cells	5000	n.a.	Actual	1	Litton Bio.	1984
Acute Inhalation Toxicity	5000	2,500 to 7,5	Estimate	n.a.	EPA Study	1980
Eukaryotic Tests (Yeast)	5000	1,000 to 10,	Survey	3	EPA Study	1980
Unscheduled DNA Synthesis	5100	n.a.	Actual	1	Bioassay	1984
Chromosome Aberrations in CHO	5200	n.a.	Actual	1	Litton Bio.	1984
Unscheduled DNA Synthesis, Rat	5200	n.a.	Actual	1	Micro	1984
Unscheduled DNA Synthesis—HeLa cells	5200	n.a.	Survey	3	NAS, Tab. 17	1982
Transformation, BALB-C/3T3	5250	$5,000 to $5	Actual	1	Litton Bio.	1984
Micronucleus Rodent Aberrations	5400	$3,000 to $7	Actual	1	Micro	1984
Acute Delayed Neurotoxicity	5400	2,700 to 8,1	Estimate	n.a.	EPA Study	1980
Acute Delayed Neurotoxicity	5400	1,000 to 12,	Survey	7	EPA Study	1980
Transformation—C3H-10T 1/2	5400	n.a.	Survey	4	NAS. Tab. 17	1982
Unscheduled DNA Synthesis, Rat	5500	n.a.	Actual	1	SRI	1984
Micronucleus Rodent Aberrations	5625	$4,350 to $6	Actual	1	Litton Bio.	1984
Chromosome Aberrations in CHO	5700	n.a.	Actual	1	Micro	1984
SCE—CHO	6000	n.a.	Actual	1	SRI	1984
Transformation C3H10T1/2	6000	n.a.	Actual	1	SRI	1984
Transformation C3H10T1/2	6150	$4,300 to $8	Actual	1	Hazelton	1984
Mouse Lymphoma	6250	n.a.	Actual	1	Micro	1984
Mouse Lymphoma	6250	n.a.	Actual	1	Pharmacon	1984
Transformation C3H10T1/2	6375	$5,200 to $7	Actual	1	Litton Bio.	1984
CHO-HGPRT	6500	n.a.	Actual	1	Bioassay	1984
HGPRT-CHO cells, V79 cells	6500	n.a.	Survey	4	NAS, Tab. 17	1982
Acute Inhalation Toxicity	6600	700 to 15,00	Survey	9	EPA Study	1980
CHO-HGPRT	6800	n.a.	Actual	1	Pharmacon	1984
Chromosome Aberrations in CHO	7000	n.a.	Actual	1	SRI	1984
Micronucleus Rodent Aberrations	7000	n.a.	Actual	1	SRI	1984
Insect Mutagenicity (Drosophila)	7000	4,000 to 25,	Survey	4	EPA Study	1980
Chromosomal aberrations—rat liver cells	7500	n.a.	Survey	3	NAS, Tab. 17	1982
Mammalian Cell Transformation test	7762	4,500 to 12,	Survey	4	van der Lee	1983
Mouse Lymphoma w/ Indep. Repeat	8000	n.a.	Actual	1	Litton Bio.	1984
CHO-HGPRT	9000	n.a.	Actual	1	Micro	1984
Mutagenicity, bone marrow test	9100	4,500 to 16,	Survey	5	van der Lee	1983
Micronucleus Rodent Aberrations	9450	$4,900 to $1	Actual	1	Bioassay	1984
CHO-HGPRT w/ Indep. Repeat	10000	n.a.	Actual	1	Litton Bio.	1984

Test	Cost/Price	Range	Data Type	Resp	Source	Date
Drosophila melanogaster	10000	n.a.	Survey	1	NAS, Tab. 17	1982
Mouse Lymphoma	11225	n.a.	Actual	1	Micro	1984
Sperm morphology—mouse	11400	n.a.	Survey	1	NAS, Tab. 17	1982
Acute Toxicity, inhalation (rat)	11613	$4,950 to $2	Survey	4	van der Lee	1983
CHO-HGPRT	12500	n.a.	Actual	1	Bioassay	1984
Cytogenetics (bone marrow)	13000	8,000 to 20,	Survey	6	EPA Study	1980
Drysophila SLRL	13500	$11,500 to $	Actual	1	Hazelton	1984
Acute Rodent Aberrations	13500	n.a.	Actual	1	SRI	1984
Sub-acute toxicity, oral (rat)	13729	10,000 to 15	Survey	7	van der Lee	1983
Insect Chromosomal Damage (Drosophila)	14000	4,000 to 30,	Survey	5	EPA Study	1980
Drysophila SLRL	15000	$12,000 to $	Actual	1	Bioassay	1984
Dominant Lethal Test (mammals)	15000	2,500 to 26,	Survey	6	EPA Study	1980
Acute Rodent Aberrations	15500	n.a.	Actual	1	Litton Bio.	1984
Acute Rodent Aberrations	16000	n.a.	Actual	1	Bioassay	1984
Acute Rodent Aberrations	16000	n.a.	Actual	1	Hazelton	1984
Cytogenic Analysis in Rodent Bone Marrow	16000	n.a.	Actual	1	Litton Bio.	1984
Acute Rodent Aberrations	18000	n.a.	Actual	1	Micro	1984
Rat Dominant Lethal Assay	18400	n.a.	Actual	1	Litton Bio.	1984
Sub-acute toxicity, dermal (rat)	18800	12,000 to 28	Survey	7	van der Lee	1983
General Metabolism Test Stds.	20000	4,000 to 50,	Survey	4	EPA Study	1980
General Metabolism Test Stds.	20000	4,000 to 50,	Survey	4	EPA Study	1980
Subchronic Neurotoxicity	20200	10,100 to 30	Estimate	n.a.	EPA Study	1980
Drosophila Melanogaster	20850	$18,000 to $	Actual	1	Litton Bio.	1984
Subchronic Neurotoxicity	21000	10,000 to 35	Survey	5	EPA Study	1980
Teratogenic Health Effects (rabbit)	26800	13,400 to 40	Estimate	n.a.	EPA Study	1980
Teratogenicity (rat)	26830	22,000 to 30	Survey	6	van der Lee	1983
Teratogenic Health (rat)	27075	13,400 to 40	Estimate	n.a.	EPA Study	1980
Heritable Translocation Test (mammals)	30000	20,000 to 62	Survey	5	EPA Study	1980
Sub-acute toxicity, inhalation (rat)	35000	18,000 to 68	Survey	5	van der Lee	1983
Mammalian Mutagenicity (specific Locus)	40000	4,000 to 50,	Survey	5	EPA Study	1980
Sub-chronic toxicity, oral (rat)	52106	26,500 to 80	Survey	7	van der Lee	1983
Teratogenic Health Effects (rats & rabbits)	53000	20,000 to 10	Survey	18	EPA Study	1980
Teratogenic Health Effects (rats & rabbits)	54000	27,000 to 81	Estimate	n.a.	EPA Study	1980
General Metabolism Test Stds.	55000	27,500 to 82	Estimate	n.a.	EPA Study	1980
General Metabolism Test Stds.	55000	27,500 to 82	Estimate	n.a.	EPA Study	1980
Subchronic Oral Dosing (rat)	56051	28,026 to 84	Estimate	n.a.	EPA Study	1980
Subchronic Inhalation Toxicity	60000	10,000 to 15	Survey	9	EPA Study	1980
Sub-chronic toxicity, dermal (rat)	62900	28,300 to 90	Survey	7	van der Lee	1983
Subchronic Inhalation Toxicity	64000	32,000 to 96	Estimate	n.a.	EPA Study	1980
Sub-chronic toxicity, inhalation (rat)	98760	48,000 to 15	Survey	15	van der Lee	1983
Reproductive Effects Test Std.	100000	30,000 to 25	Survey	10	EPA Study	1980
Subchronic Oral Dosing (dog)	103935	51,967 to 15	Estimate	n.a.	EPA Study	1980
Reproductive Effects Test Std.	114300	57,000 to 17	Estimate	n.a.	EPA Study	1980
Subchronic Oral Dosing (rats & dogs)	122000	97,000 to 18	Survey	7	EPA Study	1980
Subchronic Oral Dosing (rats & dogs)	162000	81,000 to 24	Estimate	n.a.	EPA Study	1980
Chronic Toxicity (dog)	190000	95,000 to 28	Estimate	n.a.	EPA Study	1980
Chronic Toxicity (dog)	193000	160,000 to 2	Survey	2	EPA Study	1980
Chronic Toxicity (rat)	270000	50,000 to 60	Survey	11	EPA Study	1980
Oncogenic Effects Test Std. (mice)	277800	139,000 to 4	Estimate	n.a.	EPA Study	1980
Oncogenic Effects Test Std. (rat)	377400	188,700 to 5	Estimate	n.a.	EPA Study	1980

File: OTA.COST.EST
Report: BIOASSAY COMPARISON

Test	Cost/Price	Range	Data Type	Resp	Source	Date
Chronic Toxicity (rats & dogs)	417000	350,000 to 4	Survey	2	EPA Study	1980
Oncogenic Effects Test Std. (mice & rats)	445000	145,000 to 8	Survey	8	EPA Study	1980
Chronic Toxicity (rat)	495000	247,000 to 7	Estimate	n.a.	EPA Study	1980
Whole-animal, two-species rodent bioassay	500000	n.a.	Survey	1	NAS, Tab.17	1982
Oncogenic Effects Test Std. (mice & rats)	655000	328,000 to 9	Estimate	n.a.	EPA Study	1980
Chronic Toxicity (rats & dogs)	716000	350,000 to 1	Estimate	n.a.	EPA Study	1980
Combined Chronic Toxicity (rat, mice, dog)	1064000	532,000 to 1	Estimate	n.a.	EPA Study	1980
Mouse Somatic Cell Mutation	t.b.d.	n.a.	Actual	1	Litton Bio.	1984

NOTES TO ACCOMPANY TABLE: BIOASSAY COMPARISONS

Key to Sources of Data:

Litton Bio. = Litton Bionetics, Inc.
Bioassay = Bioassay Systems Corporation
Hazelton = Hazelton Biotechnical Corporation
SRI = SRI International
Pharmacon = Parmakon Research International
NAS., Tab.1 = Toxicity Testing, National Research Council, 1984, Appendix E, Table 1
van der Lee = C. G. van der Lee, "A Comparison of Test Costs and Reporting Times by 25 Contract Labs for Tox. Research."
EPA Study = Enviro Control, Inc., "Cost Analysis Methodology and Protocol Estimates," study done under contract to EPA.

Dates:

Litton Bionetics prices were in effect until October 1984.

EPA Study was published in draft form (final version was never formally published) in 1980 — survey likely was conducted just prior to 1980.

Other prices were for years indicated.

van der Lee Study reported prices in Dutch Guilders. Data reflect only his survey of U.S. laboratories and are reconverted to U.S. dollars at exchange rate given in van der Lee study of 3:1.

Explanation of Data Types:

Actual data were from price lists from the labs.

EPA Estimates were protocol estimates constructed by Enviro Control in their detailed analysis.

EPA Survey data are mean values of the ranges.

Ranges:

Generally reflects the high and low prices from surveys.

A range given with actual data from price lists is for specific variations in the testing procedure available.

ALTERNATIVES TO ANIMAL USE IN TESTING AND EXPERIMENTATION

SECTION II: RELIABILITY

In this section, we deal with the reliability of short-term in vitro tests as accurate predictors of a chemical's in vivo carcinogenicity. The primary emphasis is placed on the Ames *Salmonella typhimurium* microsome assay, both alone and in battery with other mutation tests.

Two inherent problems exist in applying short-term in vitro tests as predictors of human carcinogenicity (13). First, the tests are conducted with non-human organisms, and second, the test organisms are actually detecting a genotoxic effect rather than carcinogenicity. Determination of reliability of short-term in vitro tests is thus dependent upon two foundation blocks:

o	Validation:	The process whereby scientists understand how the test mimics an indicated in vivo condition.
o	Correlation:	The accuracy of the predicted result, recognizing the significant degree of extrapolation necessary to go from short-term in vitro tests to human carcinoma.

VALIDATION AND CORRELATION

Validation of a test procedure may be considered in four parts:

o	Defined End Points:	The test results in attainment of definitive landmarks; not just a "change."
o	Repeatability:	Both intra- and inter-laboratory.

o	Ability to Accurately Predict:	A repeatable dose-response relationship can be demonstrated.
o	Statistical Analysis:	The test is designed to permit rigorous mathematical or statistical analysis.

Correlation among mutagenicity and carcinogenicity tests is typically expressed using a series of terms which describe the predictive power of the test (10,36).

o	Test System Sensitivity:	The number of true positives correctly identified by the model test.
o	Test System Specificity:	The number of true negative compounds correctly identified by the model test.
o	Predictive Value:	The proportion of positive responses generated by the model system which are true positive.
o	False Positive:	The number of noncorrelated positive responses in the model system.
o	False Negative:	The number of noncorrelated negative responses.

It is important to remember that the process of correlation makes a critical assumption: that a set of true responses exist, against which the model system is compared. In the case of carcinogenicity, these "true responses" are usually the rodent bioassay, either the rat or the mouse, which are considered the baseline case. This assumption has been demonstrated to be questionable, as a number of human carcinogens have not been detected in the standard rodent bioassay (10,42,44) and both the mouse and rat bioassay have significant predictive error between them (10).

ASSAY PERFORMANCE OF THE AMES SALMONELLA REVERSE-MUTATION ASSAY

The Ames Salmonella/microsome assay has been performed thousands of times in over 2,000 laboratories throughout the world and has provided results on over 1,000 chemical substances during the past 14 years. Portions of this large body of analytical data have been reviewed in over a dozen evaluation studies with the intent of determining the performance of the test as a predictor of carcinogenicity (4,6,8,30). Table II-1 presents the results of these evaluations and shows sensitivities from approximately 50% to 93%. Predictive value indices range from 58% to 100% for the chemicals in the individual test samples. The evaluation program of the U.S. Environmental Protection Agency's Gene-Tox Program represents the most recent comprehensive study of the Ames Salmonella assay. This program carefully screened hundreds of Ames reports and found a sensitivity of 78% among a sample of 123 chemicals.

In a critical analysis of each of these studies and their wide range of results, Brusick (8) has identified multiple causes for bias of the performance numbers, both up and down. These include methods of chemical selection, sample coding, use of a high proportion of chemicals known to work well or poorly with Ames testing, and differences in metabolic activation during the test procedure. Brusick concludes that a reasonably careful application of the Ames technique to a non-biased group of chemicals would be expected to yield a predictive accuracy of approximately 80% for mouse and rat carcinogens.

TABLE II-1

ASSAY PERFORMANCE OF THE AMES SALMONELLA REVERSE-MUTATION ASSAY*

Ames Assay Assessment	Year Reported	Chemicals in the Sample	Sensitivity	Specificity	Productive Value	Basis of Chemical Selection
McCann et al.	1975	300	90%[c]	87%[c]	92%	investigator selected
Sugimura et al.	1976	241	93%[c]	77%[c]	87%	investigator selected
Purchase et al.	1978	120	91%[c]	93%[c]	93%	investigator selected
Simmon, V.	1979	92	64%	78%	90%	committee selected
Anderson et al.	1979	221	90%	100%	100%	investigator selected
Rinkus and Legator	1979	271	77%[d]	NA[e]	---	literature data base
Bartsch et al.	1980	89	76%[c]	57%[c]	95%	investigator selected
Bridges et al.	1981	42	45-70%	60-80%	---	investigator selected
EPA GeneTox[a]	1984	123	78%	---	---	literature data base
NIEHS/NTP	1982	218	54%	---	76%	selection committee
Waters, et al.	1982	17	65%	NA	100%	animal germ cell mutagens
ICPEMC	1982	51	64%	35%	58%	selection committee
NCI Comparative[b]	1982	58	78%	67%	81%	investigator selected
Brusick (rat)	1983	60	78%	58%	69%	investigator selected
Brusick (mouse)	1983	60	85%	58%	69%	investigator selected

---

* All values reported are approximate and based upon evaluations of results by the authors of the reports or individuals giving their personal communications.

[a] From A. Auletta, EPA GeneTox Coordinating Committee.
[b] V. C. Dunkel, personal communication.
[c] From I. F. H. Purchase.
[d] Reanalysis of the report by B. Ames and J. McCann showed an 82% correlation.
[e] NA means not applicable.

Source: Brusick (8) with updates.

The Ames test tends to have a relatively high number of false positives in comparison with rodent carcinogenicity assays, resulting in specificity values of approximately 60% (8). It should be noted, however, that there is very little comparable in vivo negative data, and many authorities feel that the information is insufficient to make any statement of average specificity values for the Ames Salmonella/mutagenicity assay (4,54).

PERFORMANCE OF THE AMES SALMONELLA ASSAY IN BATTERY

The predictive value of the Ames Salmonella assay can be improved by combining it with additional short-term assays to form a test battery. Although no U.S. regulatory agency has yet recommended a specific battery of short-term mutation tests, most authorities recommend that an appropriate battery should include information from a minimum of three types of tests:

- Gene Mutation (Ames Test, Mouse Lymphona Test, etc.)
- Chromosomal Mutation (in vivo CHO Cytogenetics)
- DNA Damage (Sister Chromatid Exchange, Unscheduled DNA Repair Synthesis, etc.)

In addition, at least one of these tests should include a mammalian test subject used in an in vitro cell, tissue, or organ culture-assay (4).

The National Toxicology Program (NTP) currently includes seven in vitro tests within four categories in its routine evaluations for the NTP genetic toxicity program (36):

- Mutation (Salmonella, mouse lymphona)
- Chromosome Aberation and Sister Chromatid Exchange (Chinese hamster ovary)
- Unscheduled DNA Synthesis (Rat hepatocyte)
- Cell Transformation (BALB/C 3T3, SHE/SA7)

Inclusion of additional short-term in vitro tests provides enhanced predictive capability of human carcinogenicity by covering some areas of evaluation where the Ames Salmonella is not powerful. In a recent evaluation of 14 chemicals known as carcinogens in man, the Ames test alone correctly predicted 10 of the chemicals (71%). With the addition of two additional short-term tests, cell transformation and cytogenetic analysis, the four remaining false negatives were identified yielding a 100% sensitivity for the three test battery (10).

In a comparison of the predictive accuracy of various short-term tests for carcinogenicity, 18 Salmonella tests averaged 66% "accuracy" (number of chemicals correctly identified/number of chemicals tested). Comparative results from six batteries of short-term tests which included Salmonella increased the accuracy to 82-90% (23,34).

Brusick has estimated that a series of assays selected from the test categories of cell mutation, chromosome mutation and cell transformation, when used with appropriate exogenous metabolic activation, will correctly predict in excess of 95% of rodent carcinogens and essentially all chemicals identified as known or presumptive human carcinogens (10).

CORRELATION COMPARISON OF THE RODENT AND AMES BACTERIAL TESTS

The concept of correlation implies that there is a base set of data which represent a set of "true" responses to a predicted condition. In the case of human carcinogens, this set of "true" responses is considered to be the rodent bioassay.

There is a growing awareness that a number of known human carcinogens have not been detected by the standard rodent bioassay (42). Salsbury, for example, claims that the standard rodent bioassay may detect known human carcinogens no better than 50% of the time (44,45). 2-naphthylamine is one example of a human bladder carcinogen which is not identified in the rodent bioassay (42).

The predictive capabilities of the rat bioassay, the mouse bioassay, and the Ames Salmonella test were compared based on a sample of 60 chemicals from NCI bioassays. This sample showed only 72% agreement between mouse and rat, a 69% agreement between rat and Ames test, a 71% agreement between mouse and Ames test, and 74% agreement between Ames and the rat plus mouse combined. The 30% non-agreement indicates that neither rodent species is better at predicting carcinogenicity that the Ames test for the sample of chemicals (10). A comparison of the respective correlation levels for this sample of 60 chemicals is shown in Table II-2. The Ames test has model sensitivity of 78-85% compared to the rodents but predictive values of only 64-69% due to the relatively higher number of false positives in the Ames results.

In a further evaluation, 14 chemical agents established as human carcinogens by the International Agency for Research on Cancer (IARC) were compared using results from the rat bioassay, the mouse bioassay, and the Ames test. The results (Tables II-3, II-4) of this limited comparison show that none of the three model sysems are more accurate than the other. It also suggests that when used alone, none of the three models is adequate to protect man from carcinogens (10). Combining the positive responses of the mouse and the rat raises the sensitivity of the model to 86%, but this can also be achieved by combining the Ames test with either the rat or the mouse. In this study, the rat and mouse bioassays and the Ames test, when used alone, have a model sensitivity of only about 70% for known human carcinogens. Combinations of tests (rat + Ames, mouse + Ames, rat + mouse, Ames + two other in vitro tests) will raise the model sensitivity to better than 85%.

TABLE II-2

Correlation Levels Between Species Assuming Either the Rat or the Mouse as the Best Standard for Detecting Human Carcinogens.

Presumed True Standard[a]	Proposed Predictive Model	Model Assay Sensitivity	Model Assay Specificity	Model Assay Predictive Value	False Positives	False Negatives
Rat Species	Mouse Species	70%	75%	75%	7	9
Rat Species	Ames Test	78%	58%	69%	11	7
Mouse Species	Rat Species	75%	70%	70%	9	7
Mouse Species	Ames Test	85%	58%	64%	13	4
Rat/Mouse Combined*	Ames Test	78%	---	81%	---	10

[a] Species assumed to identify carcinogens and noncarcinogens most similar to humans.

* Specificity cannot be calculated for this comparison because rodent negatives cannot be true responses as some chemicals have both a "+" and "-" rodent response, and this approach uses the "+" response considering the negative response not a true indication of carcinogenicity. False positives cannot be calculated using this approach since all rodent positives are considered true carcinogens.

Source: Brusick (10) 1983

TABLE II-3

Chemicals Evaluated as Carcinogens for Humans and Their Responses In Rodent and Bacterial Predictive Assays

Chemical	Human	Rat Bioassay	Mouse Bioassay	Ames Test
4-Aminobiphenyl	+	+	+	+
Arsenic	+	-	-	-
Asbestos	+	+	+	-
Benzene	+	-	-	-
Benzidine	+	+	+	+
Bis(chloromethyl)ether	+	+	+	+
Chromium and some chromium compounds	+	+	-	+
Cyclophosphamide	+	+	+	+
Diethylstilbestrol	+	+	+	-
Melphalan	+	+	+	+
Mustard Gas	+	No Data	+	+
2-Naphthylamine	+	-	+	+
Soot, tars	+	-	+	+
Vinyl chloride	+	+	+	+

Source for:

- The Rodent Bioassay Responses

  Handbook of Identified Carcinogens and Noncarcinogens, Vol. I and II, J. V. Soderman, ed., CRC Press, Boca Raton, Florida, 1982.

- The Chemical Selection and Mutation Results

  Bartsch H., Tomatis, L. and Malaveille, C.: Mutagenicity and Carcinogenicity of Environmental Chemicals, Regulatory Toxicology and Pharmacology, 2:94-105, 1982.

Source: (10)

TABLE II-4

Accuracy of Rodent and Bacterial Predictive Tests to Identify Human Carcinogens

Standard	Model	Modern Sensitivity	False Negatives
Human	Rat	69% (9/13)*	4
Human	Mouse	79% (11/14)	3
Human	Ames Test	71% (10/14)	4
Human	Combined Mouse/Rat	86% (12/14)	2

* Data not available for one compound.

Source: (10)

In summary, the short-term in vitro tests identify less than 100% of potential rodent and human carcinogens, but standard rodent bioassays do no better and, in fact, may be less accurate than appropriately chosen in vitro test batteries.

AMES SALMONELLA ASSAY - LIMITING FACTORS

The extensive test history with the Ames *Salmonella* test has identified several factors which affect the acceptability of the assay or individual test results when any large number of chemicals is subject to testing.

Chemical Classes

The Ames *Salmonella* test has been found to predict poorly for three broad classes of chemicals: metals, halogenated aromatics, and acrylates (41). Inclusion of large numbers of these chemicals in a group to be evaluated will result in abnormally low sensitivity values for the Ames *Salmonella* test.

Potency Data

A major factor limiting the wider adoption of Ames *Salmonella* assays by regulatory agencies is the lack of definitive potency data. The Ames test is widely regarded as giving a Yes/No indicator with little or no information on the degree of change associated with increasing dose concentrations or repetitive dosing. For example, EPA's evaluation of existing chemicals under TSCA requires a balancing of reasonable risk with economic benefit. The hazard evaluation process and subsequent risk assessment rely heavily on dose response and potency information to determine what constitutes "reasonable" risk. Mutagenicity alone as an end point does not provide sufficient data to evaluate the full spectrum of risk associated with an existing chemical.

In a recent review of Ames test data from the published literature, McCann et. al (31) applied statistical methods to previously published *Salmonella* test data in a effort to evaluate dose-response relationships and potency. Using data from 152 chemicals, the vast majority of mutagens were found to produce linear dose-response curves in the initial region. Although these findings suggest that potency can be estimated from the linear portion of the dose-response curve, McCann cautions that there is a high degree of variability in the data due to differences in experimental protocol and possible toxicity effects and their findings to not definitively preclude non-linearity.

SUMMARY

The Ames *Salmonella* test can be expected to predict rodent carcinogens with a sensitivity of approximately 80%. When combined into an appropriate battery of other short-term in vitro tests, sensitivity increases to approximately 90%. The Ames *Salmonella* test tends to overpredict false positives and is relatively poor at detecting metals, halogenated aromatics, and acrylates. Traditional rodent bioassays do not predict all known human carcinogens and some comparative studies suggest that the Ames test has predictive capabilities equal to or even greater than mouse or rat bioassays. A lack of true negatives in the historical data base clouds the specificity values for the test.

The Ames *Salmonella* test ia an established assay for chemical screening and a first step in tier testing protocols. There continues to be substantial reluctance among both scientists and regulators to accept the Ames test or a battery of short-term in vitro tests as an end point for determination of carcinogenicity without confirmatory bioassays. The absence of definitive potency data is another factor limiting the acceptability of short-term in vitro test results by the regulatory community.

There are no statutory limitations to wider application of short-term in vitro tests as screening tools. However, a major policy decision is implicit in the decision as to how to treat the results of any widespread chemical screening program based solely on short-term in vitro tests.

ALTERNATIVES TO ANIMAL USE IN TESTING AND EXPERIMENTATION: ECONOMIC ISSUES

SECTION III: IMPACT OF TECHNOLOGICAL CHANGE

ECONOMIC CONSEQUENCES

Overview of the Industry

The chemical industry (Standard Industrial Classification (SIC) Code 28 in 1982 had gross shipments of over $170 billion and employed 866,000 people. As Table III-1 illustrates, this represents 8.7 percent of the total for all industry shipments in the U.S. and 4.5 percent of all employees. (This disparity in relative shares of the total for shipments and employees indicates that the industry as a whole is capital intensive and that shipments per employee are well above the U.S. all-industry average.)

Forty-five percent of the shipments and forty-eight percent of the employees in the chemical industry are for the following industries: drugs (SIC 283); soaps, toilet goods, cosmetics (SIC 284); industrial organic chemicals not elsewhere classified (SIC 2869); and agricultural chemicals, not elsewhere classified (SIC 2879).

These four parts of the chemical industry are the manufacturing sectors that are most directly affected by changes in regulatory policy concerning toxicity testing for new substances. The demand for toxicity testing services is derived from a combination of economic conditions affecting these sectors of the chemical industry, coupled with the research and development expenditures of these sectors, the technological changes in the industry permitting a host of new substances to be developed, and regulatory policy of the Federal, state, and local agencies.

The toxicity testing industry in the U.S. is estimated to have an annual value of about $650 million (47). Sixty-five percent of the testing is done by corporations in-house. The remaining 35% (about $225 million annually) is conducted by commercial laboratories, universities, and other organizations. Although there are over 110 laboratories in the U.S. that sell these testing services, most specialize in a small number of tests and are not "full service." Hazelton is the largest of the full service laboratories, with domestic sales of $36 million in 1983. Excluding Hazelton and several other large commercial laboratories, the industry is characterized by its lack of concentration, with the many small commercial firms accounting for approximately two-thirds of the value of domestic sales.

Relationship of Research and Development to Testing

Corporate research and development in the chemical industry is large and concentrated in the industrial chemicals and drug sectors. It totaled $49 billion in 1981 (35). The in-house portion of the toxicological testing industry is a part of this R&D figure. It is not possible to separate testing from the overall R&D expenditures as these functions are frequently done in the same facilities. A very general "rule-of-thumb" is that the toxicological testing industry accounts for just under 10% of the R&D expenditures in the chemical industry (47). As R&D grows, so should the testing service industry.

Projects call for healthy growth of corporate R&D in the U.S. The past few years have shown a turnaround from the overall declines in R&D, particularly in the early 1970's. A stable, growing economy with low inflation accounts for most of the growth. In addition, there have been changes in the tax laws that have stimulated R&D and capital formation. In real dollar terms, as shown in Table III-1, the chemical industry has increased its corporate R&D investment a total of 47% between 1971 and 1981, with the drug sector increasing by an impressive 70%. In comparison, R&D for all industries combined in the U.S. grew 39%. Between 1980 and 1981 alone, the chemical industry increased its R&D by 15%, about the same as the total for U.S. industry.

Table III-2 points up the very interesting fact that in 1971, Federal government expenditures for R&D were 42% of all R&D expenditures. In 1981, Federal R&D had fallen to 32% of the total, or another way of expressing it is that in 1981, corporate R&D was over 2.1 times Federal R&D. In 1981, corporate R&D was 12.9 times Federal R&D in chemicals while in 1971, this ratio was 8.9. Not only has private corporate R&D in chemicals grown, but it has expanded far more rapidly than all industry R&D and has done so while Federal expenditures in the same industry have remained constant in real dollars. In fact, the drug industry has seen a huge decrease in Federal R&D, while corporate research in drugs and medicines have grown over 70% over the ten year period.

Toxicological Testing Services

The above increase in R&D should have meant a significant expansion in the toxicological testing industry. The industry itself grew in the 1970's in response to Federal regulatory changes (TSCA, for example, was passed in 1976). But enforcement regulations were delayed, the economy slowed with recessions and rapid inflation in the late 1970's, and the growth of the industry was slowed. The companies, though, had geared up for expansion and were faced in the early 1980's with significant overcapacity--a situation which continues to exist today. This has led to prices holding fairly level over the past few years and, in some cases, price cutting to maintain market position. Even though more substances may be tested today, total revenues have remained relatively level.

TABLE III-1

PROFILE OF SELECTED CHEMICAL INDUSTRIES

SIC Code	Industry	($ Millions) Value of Shipments	(000) Number of Employees
	Total U.S.	$1,961,793	19,111
28	All Chemicals	170,085	866
283	Drugs	24,707	166
284	Soaps, Toilet Goods	25,647	125
2869	Ind. Organic Chemicals, nec.	20,007	111
2879	Agricultural Chemicals, nec.	5,255	16
	Total	75,616	418
As a percent of all chemicals		45%	48%
As a percent of U.S.		4%	2%

Source: 1982 Census of Manufacturers, Preliminary Estimates

TABLE III-2

SELECTED U.S. RESEARCH AND DEVELOPMENT EXPENDITURES

(Constant 1972$)

$ in Millions

Industry	1971	1981	% Change
Total R & D	$ 19,081	$ 26,511	39%
Chemicals	1,908	2,724	43%
Industrial Chemicals	1,051	1,306	24%
Drugs, Medicines	857	1,410	65%
Federal R & D	7,984	8,423	5.5%
Chemicals	192	196	0.5%
Industrial Chemicals	166	188	13.3%
Drugs, Medicines	26	10	-61.0%
Corporate R & D	11,097	18,087	63%
Chemicals	1,716	2,528	47%
Industrial Chemicals	885	1,118	26%
Drugs, Medicines	831	1,410	70%

Source: 1982 Science Indicators NSB Publication 83-1

Projections for the future of the industry depend on assumptions regarding a number of factors, including: growth of R&D in the chemical industry and the product of the R&D (i.e., rate of discovery of new substances), regulatory enforcement policies, both domestic and foreign, the growth of the overall economy, changes in tax policy affecting investment and capital formation, the corporate position with regard to in-house testing and contracting out the testing, and new developments in the technology of short-term tests versus traditional chronic in vivo tests.

We will not try to anticipate all of the possible changes that could occur in this analysis; however, several points are appropriate. First, international developments are important. Switzerland, for example, is considering a law banning all animal testing. As a number of companies have facilities in Switzerland, such a change could shift the sites of the laboratory testing now going on there to another county. From the van der Lee study, it is known that foreign laboratories have a significant price advantage over U.S. labs (55). A major cause of this is the strong value of the dollar in recent years. If the dollar remains strong relative to European currencies, then it is unlikely that U.S. testing labs will receive a sizable share of the foreign market. But, regulations (and prices) abroad can and will affect the U.S. testing services.

The other significant factor is U.S. regulatory policy. As mentioned above in Section I, if, in the future, data bases and technologies are developed for short-term in vitro tests that show the reliability to be equal or close to the results for traditional tests, and, if regulatory agencies respond with approvals for using these tests, not only as a screen but as conclusive, then a very rapid change in the demand for and composition of testing servces will occur. And, with this change, it is probable that new and cheaper methods of conducting the tests could be developed that will radically change the industry as we now know it. It could mean that the total value of testing services as measured by expenditures and sales (not by social benefits and costs) will remain equal or be smaller that at present, with many more substances being tested.

An unknown in the above analysis is the amount of testing on substances aimed at the U.S. market that is done in the U.S. labs and the amount done abroad. Likewise, it is not known how much testing that is done in U.S. laboratories is for meeting regulations of foreign regulatory agencies. A change in this balance could mean significant changes in the domestic testing industry. A special survey of laboratories would have to be conducted to determine this effect--we know of no such data currently available.

Implications for the U.S. Budget

The programs within the U.S. Budget that are identified with toxicological testing are: The Environmental Protection Agency's Division of Toxic Substances, the National Center for Toxicological Research in the Food and Drug Administration (part of the Department of Health and Human Services), and the National Institutes of Environmental Health Sciences' Applied Toxicological Research and Testing Program. Other programs are not identified with separate budget line items and are assumed to be scattered among the various agencies and departments.

The 1985 Federal Budget (1) lists the following data concerning the above programs:

		($000)	
	1983	1984e	1985e
EPA			
Toxic Substances (S&E)	33,306	33,780	38,327
Toxic Substances (R&D)	14,166	12,352	24,863
Toxic Substances (Abatement, Control, Compliance)	22,506	22,650	26,451

FDA			
National Center for Toxicological Research	19,614	21,555	21,575
HHS			
National Institute of Environmental Health Sciences	165,498	183,384	284,405
Applied Toxicological Research and Testing	52,258	55,257	55,054

Adding together the toxicological research program budgets (excluding the enforcement function), the fiscal year 1985 outlays will equal $91.5 million. Although a small part of the entire Federal budget (.01% of $925 billion total outlays), it still represents a significant amount (14%) of the estimated $650 million domestic toxicological testing industry. A change in direction of Federal R&D in toxicological testing could definitely have an effect on the industry. (Note: This argument concentrates on policy decisions affecting the direction of research, not on regulatory policy changes.) From our discussions with Federal personnel, we do not know of any major policy changes that are anticipated. But programs are being conducted in developing sizable data based on short-term tests by the EPA, and there is a program at the Oak Ridge National Laboratory on short-term testing for mutagenicity. In addition, there have been Federally sponsored conferences and workshops that have stimulated research in short-term alternatives to chronic in vivo testing for mutagenicity and carcinogenicity.

Federal R&D expenditures for toxicological testing has not shown a definite trend over the past few years. The programs have remained relatively constant in funding. Some cases such as the HHS Applied Toxicological Research and Testing Program have even seen a very slight decrease. Accounting for inflation, this has effectively meant a real purchasing power decrease of five to six percent over the last two years. The EPA toxic substances R&D budget took a 13% cut from 1983 to 1984 but had its funds restored to the 1983 level for fiscal year 1985.

These agency budgets are part of the "discretionary budget." Over the past several years, with the growth of the Federal deficit and the Administration's policy of finding areas to cut expenditures, it is noteworthy that these programs have held their own and even had some moneys restored. Countering this overall budget-cutting policy has been the Administration's supportive position on basic research. They have encouraged some expansion in basic research by increasing the budgets of the NIH, the NSF, NASA, and selected other research oriented agencies, while cutting back on the more applied and development expenditures that they feel are best done in the private sectors.

Looking ahead, the above trend will probably continue. Basic research programs will be continued at approximate level funding, with possible real purchasing power cut by inflation rather than the budget cutter's pen. Unless powerful arguments can be made through the political process for a sizable increase, these programs will continue at their current levels. It is not likely that entirely new programs will be heavily funded but that a reprogramming of existing funds will have to suffice if a new area or technological development occurs that demands Federal support.

Referring back to Table III-2, detailing R&D expenditures in the various sectors, it is interesting to note that, as the Federal research in drugs and medicines declined by over 60% in the 1971 to 1981 decade, the private corporate funded research and development has expanded by 70%. For all industries in the nation, this more rapid growth of corporate research over government research has been a long term trend. The chemical, and particularly the drug sectors, have been substituting corporate R&D for government R&D at a more rapid rate than the national average for all industries. Given current Federal policies, we expect this trend to continue.

Most government expenditures are for research. It is not known how corporate reporting of R&D is allocated between research and product development, or exactly where they account for toxicological testing expenses--as research or as development. Thus, rapid expansion of corporate R&D could mean level or decreasing testing as a percentage of the total R&D. Additional survey data would be needed to determine the answer to this empirical question.

POLICY CONSEQUENCES

The rapid increase in development of short-term in vitro tests to support the health and environmental risk assessment process raises questions concerning the implications to U.S. regulatory policy. In approaching this area of policy consequences, we shall focus upon four areas: the implication to risk assessment methodology, statutory limitations and regulatory policy constraints, public perception, and international implications.

As in the previous sections of this chapter, this discussion is limited to the Ames *Salmonella typhimuirum* microsome assay, alone and in batteries similar to that recommended by NTP and the OECD. Policy consequences are considered for the two agencies primarily concerned with regulatory policy, the Food and Drug Administration, and the Environmental Protection Agency.

Implications to Risk Assessment Methodology

The Ames microsome assay is routinely accepted as part of a screening protocol for chemical hazard analysis. Typically combined with two to four additional short-term in vitro tests, the results are usually the first testing step in a tiered approach to evaluating chemically induced mutagenicity.

Although widely utilized in hazard evaluation, the Ames test and batteries of additional short-term assays are almost universally regarded as limited to screening tools. They are not considered substitutes for traditional in vivo animal tests in a "definitive" determination of carcinogenicity. In evaluating the potential policy implications of an expanded use of short-term in vitro tests, we will consider two hypothetical situations:

- Greatly expanded application of short-term tests as screening devices
- Gradual substitution of short-term tests for traditional in vivo tests

Expanded Screening Applications:

No policy restraints are envisioned in making greater use of short-term tests, especially where their cost can be held relatively low. In its recent review of toxicity testing needs and priorities, the National Academy of Science (34) recommends development of a procedure for systematic screening of approximately 70,000 chemicals, as a basis for nominating candidates for detailed testing. One method of accomplishing this screening is development of one or more very-low-cost (VLC) tests, such as *Salmonella*/microsome assay tests, that could be broadly used. At this point, there is no concensus on exactly what VLC test or battery is appropriate, but the concept of a primary screening tool is generally accepted, so long as it is not used to define the degree of concern over a tested chemical.

A major policy issue does emerge concerning the decision on how to treat the *results* of any large scale screening program. A program screening thousands of chemicals could conceivably identify tens or hundreds of positives from the Ames assay or a battery of short-term in vitro tests. If these chemicals were designated as potentially carcinogenic, based on this screen, there would be substantial impetus to quickly conduct the long-term bioassays necesary to verify the positive determination. The numbers of chemicals identified would undoubtedly completely overwhelm the available capacity to conduct the bioassays. Given the known tendency for the Ames test to give elevated numbers of false positives and the critical lack of true negative data in the historical data base, many chemicals may be "case in doubt" that are actually non-carcinogenic. This period of uncertainty while resolution is achieved could have major cost and social impact for chemicals in widespread use in commerce.

Substitution for Animal Tests:

As shown in Section II of this chapter, the predictive accuracy of the Ames Salmonella/microsome assay as a predictor of rodent carcinogenicity is approximately 80% (8). In combination with other in vitro tests, an appropriate battery may exceed 95% specificity (10). Although the battery concept shows a high degree of correlation with rodent-determined carcinogenicity, currently, there is little support among the scientific and regulatory community for replacement of traditional animal tests with short-term in vitro test batteries.

This reluctance is based upon a number of expressed concerns:

- Microbial systems do not "intergrate" effects as well as an in vivo animal system.
- Some tests do not provide a dose-response curve which is an essential component of risk assessment methodology.
- The tests may not be responsive to certain inducers that do not interact directly with DNA.

Regulatory policy in the Environmental Protection Agency and the Food and Drug Administration has encouraged the use of short-term in vitro tests but only as part of a tiered screening approach. No standardized test battery has been agreed to by the

Agencies and attempts to define a microbial bioassay screen for inclusion in the recently published EPA FIFRA Part 158 rules on Data Requirements for Pesticide Registration (14) were rejected by agency scientists. There is widespread feeling among EPA personnel that scientific developments in short-term in vitro testing are evolving rapidly, and it is unwise for the agency to adopt the rigid constraints implicit in any defined battery of tests (40). The agency prefers to suggest a number of such tests, without defining any specific one or a single defined battery.

In summary, agency risk assessment methodology would require a change in regulatory policy in order to utilize short-term in vitro tests as replacements for existing animal tests in hazard evaluation and risk assessment. No policy changes would be necessary for expanded screening evaluation of chemical substances using short-term in vitro tests, but there would be a major policy decision implicit in what to do with the results of a wide ranging screening program.

Statutory Limitations and Regulatory Constraints

The major pieces of legislation requiring chemical product testing are:

- Toxic Substances Control Act (TSCA)
- Federal Insecticide, Fungicide and Rodenticide Act (FIFRA)
- Federal Food, Drug and Cosmetic Act

These statutes require testing to demonstrate no unreasonable adverse effect from the proposed use but do not specify specific kinds of tests or testing protocols. There is, therefore, no statutory limitation to the broader use of short-term in vitro tests in screening programs or as replacements for existing animal studies.

Limitations on replacement use of short-term in vitro tests are implicit in agency regulations (rules) which spell out the type of testing that shall be done to evaluate the introduction or expanded use of chemical substances. This may include requirements for specific tests (acute oral toxicity - rat, dermal sensitization, 90-day feeding study - rodent, gene mutation, etc.), but still allow some degree of flexibility in how certain tests are chosen, especially in the area of gene mutation. For example, the recently published 1984 FIFRA Test Rules (14) specify that the required battery of mutagenicity tests must include tests appropriate to address three categories of interest: gene mutations, structural chromosomal aberrations, and other genotoxic effects as appropriate for the test substance. Choice of the specific tests to provide these data is left to the applicant.

Finally, the agencies publish Regulatory Guidelines that suggest protocols for many tests and provide general guidance on how the agency feels a given test should be conducted. An example is the Guidelines for Acute Toxicity Testing, recently published by the EPA's Office of Toxic Substances. These guidelines do not present an impediment to wider use of short-term in vitro tests.

In summary, there are no statutory impediments to broader application of short-term in vitro tests, but there are regulatory limitations (test rules) that require continuation of traditional animal tests in some areas that have been considered as candidates for replacement tests. A change of this magnitude would be considered a policy change by EPA and would have to be approached through formal rule-making, a process that could reasonably take two or more years. More immediate flexibility is possible via an interim policy which could be promulgated by the agency administrator for immediate implementation.

It is clear that any change in policy resulting in replacement use of short-term in vitro tests would not be considered without the concurrence of the agency's outside scientific advisory and review groups. The role of EPA's Scientific Advisory Panel, for example, would be crucial in proposing an interim policy or rule-making under FIFRA.

A policy-related change typically requires intra-agency review and evaluation prior to adoption. As such, any proposed replacement of short-term in vitro tests for selected animal tests would receive such a review, thus reducing the potential for impact within related agency divisions.

Public Perception

Public awareness over increased adoption of short-term in vitro tests is potentially directed toward three questions:

- Does the proposed change provide equivalent protection of human health?
- Does it provide equivalent protection from environmental degradation?
- Does it provide for reduced use of test animals?

It is likely that the issues involved in this case are sufficiently technical that the public is not likely to embrace it with a high level of concern. From 1972 to 1976, for example, the U.S. Food and Drug Administration tested almost all substances on the GRAS List using Ames and tiered tests. This resulted in the GRAS Affirmation List, but the general public took no great note of the testing program (15). The exception is the advocates of animal welfare who have directed their attention to the potential for reduced use of animals in testing programs. It is anticipated that these groups would support increased utilization of short-term in vitro tests, if it resulted in a decreased level of animals required for testing programs.

International Implications

The United States is party to several international agreements concerning guidelines for testing of substances for use in international commerce as well as reciprocal licensing arrangements for registered food, drug, and chemical products. The United States participates as a member of the OECD Test Guides Update Panel. A major goal of the group is the concept of facilitating the acceptance of testing by OECD members so that there is mutual acceptance of data. It is current EPA policy that any testing guidelines published by the Agency will be in harmony with OECD guidelines. It should be noted that OECD promotes the concept of Mutual Acceptance of Data, but this does not extend to mutual assessment of hazard evaluation and risk assessment, which is retained by regulatory agencies of the individual countries. There are no international impediments to the U.S. increasing its utilization of short-term in vitro tests, but it is unlikely that any unilateral action would be taken which is at substantial variance with OECD's guidelines and statements of principals.

As one of the world's countries with large resources for substance testing, the United States can be expected to carefully consider the international policy implications of any change which potentially affects the movement of such products in international commerce.

Summary

Expanded use of short-term in vitro tests as a screening tool for mutagenicity and presumptive carcinogenicity has no major policy constraints. The technology, when packaged in appropriate test batteries, is currently accepted by the scientific and regulatory communities. A decision to apply initial screening to most of the 70,000 chemicals identified in the NAS toxicity analysis (34) would require a policy change due to the resources necessary to accomplish it, but the use of short-term in vitro tests as a screening device could be accomplished within existing policy. A major policy decision is implicit on how to treat the data resulting from such screening and the accompanying potential for a massive bioassay program to verify the positions resulting from such a screening.

Replacement use of short-term in vitro tests for existing animal tests would be considered a fundamental change in agency risk assessment procedures and would have to be supported by policy change. There is no statutory impediment to such a change, but changes in test rules would be required. These could be accomplished through normal rule-making procedures but would likely take several years to accomplish. Additional flexibility can be obtained via an interim policy decision by the agency administrator, but any such changes would have to be well documented, supported by inside and outside scientific review, and be reconciled with U.S. participation in international testing agreements.

ALTERNATIVES TO ANIMAL USE IN TESTING AND EXPERIMENTATION

SECTION IV: ECONOMIC ANALYSIS OF BASIC AND APPLIED RESEARCH

INTRODUCTION

This section will review the types of economic analyses that have been and could be applied to measuring the risks, benefits, and costs of using animals in basic and applied research.

This topic represents only one example of a variety of situations where economic analyses have been applied to new and rapidly emerging technologies. Since the literature on the risks and benefits of alternatives to whole-animal testing is limited, this summary will deal with the use of economics in technology assessment in general as well as with appropriate applications to alternatives to whole-animal research and testing.

The basis premise of this section is that it takes a relatively long period of time for a technology to mature to the point where it is not only accepted in the market place, but where it has diffused throughout the industry and is used routinely. At this mature stage, the production process (the combination of labor skills, materials, and plant and equipment) is settled and the state-of-the-art is used by most, if not all, producers. (But note that there is always room for an innovator to find a more efficient method of production and reap additional profits for a short period of time. If this occurred, the industry would be in a stage of renewed growth and change that occurred after it had reached a degree of stability.)

At least twenty years is often required to reach a mature stage of development. Economic measurement of the effects of a technology on the economy is very difficult until the technology becomes an integral part of the economy. Various estimating techniques are applicable to intermediate stages, but in a fast changing environment, all empirical studies are subject to great potential error and variability. For the most part, the economic studies of new technologies are case-by-case analyses for which documentation exists. It is impossible to generalize from an individual case to the general impact of all related technologies because of the tremendous differences between each application of the technology and between the various producing firms.

Economic analyses for evolving technologies are important. But they represent a planning horizon viewpoint that is short-term and not well suited for long-range policy decisions. As a series of short-run analyses are completed and as time elapses and the technology in question continues to develop, the summation of experience with many economic stuides of that technology may yield a consensus opinion for the future impact of the technology. This consensus opinion, correct or incorrect, is based upon a strongly documented foundation so that public opinion as well as economic thinking will permit the results of these economic studies to be used in policy formulation.

Alternative tests to replace the use of whole-animals in research and testing is a new technology. The Ames test, the most popular and the best documented of all the in vitro tests, is barely fifteen years old. The studies and analyses are just beginning to gather and measure some economic factors associated with these new methods. There is no standard in the industry. There is no uniform protocol among the various testing laboratories. There is no knowledge of the real risks and benefits of these tests (some would argue that there is no knowledge of the real risks and benefits of more traditional whole-animal research and testing either). And there is no consensus on the results of the studies on alternatives to whole-animal tests on the scientific level. Therefore, it is premature to base policy decisions at a national level on the outcome of economic and risk/benefit analysis. However, it is appropriate for individual businesses to use economic and risk analyses to make decisions regarding competitive market entry, pricing, and future business planning. And, it is appropriate for legislators and policy makers to consider social risk/benefit studies as one input into decisions, recognizing that the results are conditional, not determinative.

The process of technological change is discussed below, and the applicability of economic models at various stages in the development and maturation of the technology is addressed along with an assessment of the validity of using economic models for policy decisions regarding alternatives to whole-animal research and testing.

THE PROCESS OF TECHNOLOGICAL DEVELOPMENT

Tracing new technologies from their inception as an idea to their acceptance, use, and diffusion throughout the economy is a complex task. Authors that have researched this process have found as many different ways that technologies have developed as there are technologies (19,37). Very little consensus has been reached in the studies. Most authors cannot even agree on the time span it takes to get a new technology into the mainstream of use (33). A theory of innovation that embraces all situations, variables, technologies and types of impacts has yet to be developed (37).

There is good reason for this divergence of opinion. Primarily, most researchers are asking different questions, looking at technologies differently, and measuring different things (21).

Does a new technology begin with an idea, or does it begin with someone funding the inital studies (possibly in very unrelated disciplines) and applying the idea to a potential commercial end? Does first commercial use occur when it is first used by a business, profitable or not, or when it begins to make a profit for a firm. Does diffusion mean when it is copied by many firms, or just when a single firm sells a given amount or proportion of the product?

What is the role of scientific research in the process? How valuable are the results of fundamental science? Can the idea generator capture some of the eventual profits to be made from the technology? Does it matter to society whether the fundamental researcher profits or not? Does it matter to the stimulation of ideas whether there is profit in it or not?

When we study the economics of technology are we interested in measuring the benefits and costs to an economic enterprise (defined as any microeconomic unit in the market place), or are we measuring the benefit and costs of "public goods" for consumption by our society as a whole?

The answers to these (and many other questions) will determine the type of economic study to be undertaken, and to a large degree, the results. An economic model can be built for almost any situation. Its results cannot usually be applied to other situations that may be analogous because most models are for unique situations and employ a number of limiting assumptions. If such a model is then implemented through a variety of measurement techniques, then an additional set of assumptions concerning the derivation and use of statistics is overlayed on the model's assumptions. The caution cannot be too strongly emphasized that one should not generalize to all cases the results of specific technological studies. Further, the earlier in the development of a technology, the less reliable the results of economic studies are likely to be.

ECONOMIC MODELS

Uncertainty, Risk, and Benefit

An economist would define risk as a situation where the outcome is not certain, but where the probability of each possible outcome is known or can be estimated (28,43). Risk, therefore, must be distinguished from uncertainty. Risk to an economist is measurable. Uncertainty is a condition that is not measurable, even in a probabalistic framework.

Real benefits are occurrences that augment society's welfare (28). In contrast, pecuniary benefits arise from changes in relative prices as the economy adjusts to the results of a new technology or set of conditions. These pecuniary benefits may change the distribution of income among the various sectors of the economy, but gains of one party will be offset by losses of another and they are not benefits to society as a whole.

A business firm trying to increase its market share and/or profits is vitally interested in pecuniary benefits and costs. A governmental organization should be primarily interested in the real benefits and costs of its policies. This important distinction is often overlooked.

Several types of analyses have been used by economists to measure benefits, risks, and costs of many types of governmental actions, ranging from the construction of dams to the regulation of drugs. A number of laws require government agencies to assess the benefits and costs of proposed projects prior to their initiation (24).

A very commonly used type of analysis is benefit-cost. If the discounted percent value of a stream of benefits over a given time span into the future is greater than the discounted present value of the costs over the same time frame (i.e., the benefit/cost ratio is greater than one), then the project is viewed as beneficial and should be undertaken.

Another related type of analysis is cost-effectiveness. The object of this exercise is to find the least expensive method of obtaining a desired governmental goal. The goal could be expressed as a benefit to society, and the analysis of the different ways of creating that benefit are evaluated with respect to their costs. The methodology is very analogous to benefit-cost and is subject to similar methodological problems.

The third type of analysis involves estimating risks. Risk analysis takes what is known or what can be estimated with some degree of certainty and translates the many possible alternative courses of action into a framework for decision making. A desired impact or outcome may be the objective and the "best" route for getting there can be traced. Or, from a starting point such as the existence of a new technology, the various routes to the market can be traced and the optimum path for success plotted. Diagrams based on these estimates of the risks involved at each decision point on the path are called decision trees.

Risk assessment and measurement is not the same as risk management. Risk management involves the actual system used by policy makers to initiate programs and to attempt to balance social costs and social benefits. Economic analysis is, in theory, a measurement of the inputs into a system that makes the risk manager more aware of the options available. Although decisions can be made from the results of economic studies, much more is involved in merging those results with political realities.

The focus of this discussion will center on economic risk measurement and its applicability to the evaluation of new technologies, keeping in mind that the outputs of the measurement of risk frequently become the inputs to the management of risk.

The Measurement of Economic Risks and Benefits in Emerging Technologies

This section will briefly digress from the main thrust of the chapter to discuss the numerous special problems an economist or statistician will encounter when dealing with new technologies. Much has been written about the problem with surveys, statistical probabilities, secondary data sources, etc., but very little about the application of measurement techniques to new technology.

The most significant barrier to any benefit/cost, benefit/risk or other evaluation that requires measurement is obtaining good data. New technologies, as characterized above, are in a state of flux. The industry is likely to change daily. New discoveries, new ways of manufacturing the product, and new sales methods, act together to make measurement difficult. Standard national industry statistics are often several years old (the 1982 Census of Manufacturing compiled by the Census Bureau, for instance, released preliminary data in 1984). When national data sources are current, they tend to be highly aggregated and impossible to split into the fine components needed to evaluate the impact of new and often relatively small industries.

A direct survey of firms in the industry developing the new technology is possible. But even the firms themselves may not know the answers to the economic factors that would be necessary for evaluation. Sales may be combined with other data. Employees may be dividing their time between the new product and more mature products of the company, and managers may not have a good grasp of the labor inputs actually being devoted to the new technology. Prices, inventories, orders, etc. may be rapidly fluctuating, making estimates impossible.

Even more important than the above factors is the operation of the competitive market. A new and better technology is a valuable property right to a company. It is unlikely that a manager will be willing to be open about his production process, costs, marketing plans, etc. A leak of plans or secret formulae could ruin a company's market position, and most data will be held as proprietary. They will typically report only what is required by law (i.e., for securities issues, tax returns, etc.).

As a technology becomes widely spread, adopted by many firms, and sold commercially, the valuable property rights will either have been formalized in patent protection or will have been appropriated by the innovating firm in the early years of production. Then firms might be more willing to cooperate with surveys and other data gathering exercises.

In summary, it is almost impossible to gather timely and accurate data for new technologies with any degree of reliability because of confidentiality, rapid changes that make data obsolete even before they are published, and honest lack of appropriate information at the firm or source level.

Measurement in Benefit-Cost models has been a widely debated subject (28,32,24). The benefits and costs that are associated with known factors can be evaluated more easily than those associated with a large number of unknowns. New technologies are in the latter category. Many assumptions have to be made concerning the future and the likely demand curves for the product, the likely sales prices and values, the likely production process and associated input costs, etc. As has been described above, these factors are less well known and more rapidly changing for new technologies than for existing ones. For policy purposes, benefit-cost analyses applied to the early development stages of new technologies is not the most appropriate tool to use.

Added to the special problems of new technologies are the problems of assigning values to human life. Biological testing risks are those that, even more than in many other areas, involve human exposure to potentially harmful substances. Although the courts and insurance companies do assign values to human life and limb, the "true" value to society is completely unknown, and its determination often becomes a moral and ethical question.

Benefit-cost calculations do not take into consideration the economic redistribution that may occur in the future. The group in society that receives the benefit may not be the same group that pays the cost. This type of analysis is incapable of evaluating that effect.

Benefit-cost analysis often focuses on one technology. It does not evaluate alternatives or substitutes. This must be handled by doing a series of such analyses for each possible alternative to see which one may be most "cost-effective." But, there is no assurance that the evaluator has looked at all the possible alternatives. Further, in dealing with new technologies, the other possible alternatives may be on the drawing board, yet to be discovered. The time frame of the benefit-cost analysis can be a stream of benefits and costs well into the future (streams of ten to fifty years are not uncommon), yet a radically new discovery could truncate the benefits from the subject analysis long before the time frame being used is completed.

Finally, a value judgment must be made concerning the discount rate to be used into the future. There are as many possible discount rates as there are economic projections. Most macro-economic models (the type used for predicting changes in interest rates, GNP, and other aggregate variables) have reasonable accuracy a few months to a few quarters into the future. None can predict ten to fifty years with any degree of confidence. It is anybody's guess as to what the time value of money will be.

The discount rate choice is critical since it materially affects the outcome of the analysis.

Measurement of risk is more appropriate than benefit-cost analyses when the technology is closer to its formative stage and many factors can be assessed by probability functions. This permits quantitative judgments to be made on values, assumptions, and outputs. Any decision maker is able to substitute their own values. Sensitivity analyses are relatively easy to do. Such exercises will show the expected variation in results with different values assigned to different decision points. The major drawback to this type of analysis is that the results are only as good as the judgments and assumptions that went into the model, and with multiple probabilities assigned along the decision path, the likelihood of the output being realized diminishes as the probabilities get multiplied together. Or conversely, the number of possible outcomes increases dramatically.

Limitations and Philosophical Concerns

Approaching risk models from purely an economic standpoint can be misleading. Measurement is only one factor in risk assessment. There is risk in almost everything that is done, from maintaining the status quo to radical change. What is acceptable risk or unreasonable risk is a matter of policy and public opinion--it is not primarily an economic measurement problem. Many questions emerge: should government officials make these value judgments for society? If so, what criteria should be implemented? Should it be taken to the voters in a poll? In short, what do the measurements mean for policy purposes? These are political questions, well beyond the scope of the economist's or statistician's tools to evaluate (28).

At this point, a basic philosophical decision has to be make. Should decision makers use the results of these quantitative studies as guides to policy or should they resort to other methods for policy actions. There are other methods. For instance, our legal system is based on common law precedents, procedural regularities and codified statutes. Legal suits are initiated by injured parties making claims against those they believed injured them. This permits the award, in meritorious cases, of money damages to the injured party. Risks, damage estimates, and claims become the domain of the negotiation process and the jury system. As imperfect and expensive as this system may be (each separate case must be settled in an often lengthy court trial), it does result in precedents being set for damages and in an economic value being placed on the risk taken. Over time, this may result in codified and relatively set damages being awarded for specific harm. There are many drawbacks to the system. Among them are the chance of unusually radical departures from average awards (a sympathetic jury situation), the cost of starting and concluding a trial being so large that many injured do not initiate a claim, the often long time span between injury and award, the possibility of not recovering from a company no longer in business, bankrupt, or unable to pay.

ECONOMIC MODELS OF BASIC RESEARCH

Economic analysis of basic research investments and outputs presents extremely different and difficult problems from the analysis of research oriented toward marketable final products. Basic research, also often referred to as the search for fundamental knowledge, frequently has no identifiable output in the commercial world. Economic data are inadequate to quantify basic research benefits and costs for a variety of reasons, including:

- the long time frame between investment and final consumer products
- use of knowledge by firms or industries unrelated to the discovering entity
- difficulty of tracing back to original inventory/discoverer
- knowledge, once disseminated, becomes a building block for further discoveries and, therefore, does not depreciate over time, while inventions using the knowledge have an economic value that does depreciate over time.

Some attempts have been made to quantify the results of basic research. Among them are measuring the patents issued to organizations primarily conducting basic research, measuring the journal articles published by recipients of grants and by institutions conducting basic research, and by tracing on a case study basis final products back to the original research.

None of these studies are definitive. Patents are a very imperfect measure for output--many findings are not patentable under the rigorous definitions of what is a patentable invention. And some companies and industries prefer keeping trade secrets and research results hidden from public view as a better economic approach to protecting markets than taking out patents and publicly disclosing the information.

Journal articles vary greatly in quality. Over the past twenty years, the number of journals has increased dramatically in the physical and social scientists, reflecting the need for researchers wanting to advance in their careers to "publish or perish." Some researchers are far more prolific in their writings than others and the quantity of articles written on a subject is no substitute for quality or influence in the profession or on future innovations that will have an economic impact.

Most economic literature that has dealt with basic research has looked at aggregate levels of investment and the contradiction in our free market oriented society between our desire to have basic research conducted and our inability to reward the discoverer of new fundamental knowledge adequately (3,38). The contradiction hinges on the fact that the output of basic research is knowledge and that knowledge (or information) is indivisible and inappropriate. This means that once published, the knowledge cannot be divided into its parts and assigned value. What is useless to one individual may be invaluable to another. Since information is intangible, it is almost impossible for a writer, researcher or publisher to protect their rights to the information and receive payment for the information. Legal attempts to protect information exist (e.g., copyright laws), but they are difficult to enforce adequately.

Yet, information is a commodity and can generate profits. But the purchaser does not know the value until he has the information; therefore, it can usually be acquired at relatively low cost and the market place cannot put an accurate value on it a priori (3).

In theory, as described by Nelson (38), we can expect from a given expenditure on science a given flow, over time, of benefits that would not have been created had none of our resources been directed to basic research. This flow of benefits, properly discounted, may be defined as the social value of a given expenditure on basic research. Corresponding social cost is the opportunity cost (allocation of benefits which we have deprived ourselves by spending on basic research). The difference is the social profit.

But, because basic research is likely to generate externalities (economic consequences beyond the firm or industry making the investment) that are unpredictable, the conduct of much basic research is left to the public sector, or to the private section only when the costs are low and the potential returns high. Thus, only some large firms and government entities will invest in basic research. As research becomes more directed toward product specific ends, the economic motive for private investment increases.

This discussion is important only to emphasize that the motivation for basic research goes beyond measurable economic goals and that the applicability of economic analyses to the conduct of basic research is poor, given the ability to measure outputs of basic research.

Therefore, much can be said and argued about the appropriateness of using animals for basic research from the scientific viewpoint but very little from the economic view. Economic arguments will take two routes. The first is to try to estimate the expected flow of benefits based on historical experiences. What have we learned from research on animals (that could not have been learned from other methods) that has resulted in eventual commercial payoff? This type of analysis assumes that specific products can be identified and will say nothing definitive about present research and its future payoffs.

The second route that economic arguments take is the costs of using animals compared to the cost of alternatives. But for any research where the scientific community feels that animals are the only viable method, costs become far less important than results. Further, as demonstrated in Section 1, for most research involving animals, the cost of the animals (as well as other non-human labor inputs to the research) becomes relatively minor as both the time horizon and the amount of human labor increases. Most typical basic research projects involve many years of high-skilled human effort and, therefore, fall into the category where the costs of animals and physical equipment is not a determining factor.

In short, applying economic models to measure the benefits and costs of basic research in a particular narrow field or for a particular research project or grant would be unproductive. As research projects move toward identifiable market-type products, the applicability of economics increases greatly in measuring outcomes to society.

CASE STUDY: ALTERNATIVES TO WHOLE-ANIMAL TESTING

The use of short-term in vitro tests to replace whole-animal tests presents an interesting case study. This technology is unique in that it combines elements of both a new process and a new product. A process related technology is one that affects the production function--the combination of labor, capital, land and entrepreneurship that is brought together to manufacture a product. Process related technologies tend to be adopted quickly and to be spread quickly from the innovating firm to the others in the field.

Biological testing for the safety of a product for human consumption is both a prudent investment for a company and a requirement of government regulations. It is a process related technology--one that is used in the production of a consumer good and is not directly consumed by the final purchaser. (In some cases, the chemical substance itself is not sold for final consumption. An example is the agricultural chemical that the farmer uses to improve the yield of his crops.)

Over the years, an industry has developed that specializes in testing new substances. It is a service industry and the products of the industry are test results sold commercially. Usually, the purchaser uses the results as an input to their own production (in economic parlance, it is an intermediate good). From the national income accounting perspective of the economy, this industry can also be viewed separately. It combines the factors of production and sells its output commercially. In fact, not only is there research being done to develop new chemicals and substances for final consumption by the public, but there is also much research being done to develop new tests and testing procedures.

This testing industry has and will exist whether there are in vitro short-term alternative tests or not. Animal research and testing is a significant business (see Section III). What may change as a result of new technologies is the structure of the industry and the firms within it.

The social benefits that may accrue from using alternative short-term tests for mutagenicity and carcinogenicity are:

- decrease cost of testing, freeing resources for other purposes (including testing of additional chemical substances)
- rapidly increase the data base on the outcomes of these tests on chemical substances
- prevent injury or death to those people who may be exposed to substances that were not tested under status quo protocols because of time and/or resource limitations. Related benefits would include such factors as increased productivity at work, less time lost from the job, increased life span, etc.
- stimulate the development and introduction of new safe substances into the market

- reduction in the number of animals used for research and testing and more humane treatment of animals in general.

The list of benefits is based on an assumption that short-term alternatives to whole-animal tests accomplish their purpose, and there is no probability that they are incorrect or ineffective. Given today's scientific knowledge about all testing protocols (in vivo as well as in vitro), this is an unrealistic assumption. There is no evidence that animal testing is conclusive proof of carcinogenicity in humans for all substances tested, just as there is an insufficient data base for showing conclusive proof of short-term testing outcomes (refer to Section II).

Society is operating somewhere in the murky, unknown area with all of these testing procedures. Most of the benefits listed are hypothetical and of the form: if short-term tests are at least as reliable as chronic whole-animal tests, and they continue to be significantly less expensive (including costs associated with risk), then the benefits will be very significant.

The problem is not that simple. As pointed out above, the technology is still emerging and developing. We, as a society, are in a stage of great uncertainty concerning testing for carcinogenicity. There are very large risks associated with mistakes, and examples are numerous where such mistakes have caused considerable harm.

There are, therefore, two problems to the aspects of costs. One is assigning risk values to those potential costs that can be measured. The other is to discuss uncertainties that, by definition, are just too speculative to evaluate empirically.

In focusing on the societal costs associated with short-term testing, the only relevant question is: what are the additional costs or risks that society incurs by accepting short-term tests as a replacement for chronic in vivo testing?

This additional cost centers on the probability connected with false outcomes from the test results. If the probability of finding false negatives or false positives is greater with short-term alternative tests, then the margin by which it is greater becomes the basis for calculating social costs.

(The actual calculation of the costs would involve the same parameters as for whole-animal testing outcomes that might be incorrect. The considerations that would go into such calculations include: value of lives lost, value of human pain and suffering from injury, etc. for false negative findings. For false positive findings, the costs are those associated with not bringing a useful product to market and not making it available to consumers who might benefit. A large body of literature exists on the specific factors associated with risks in the biological testing area. Our purpose here is to discuss the validity of using economic analyses for policy decisions. A comprehensive review of the theoretical and measurement problems is beyond the scope of this section. Nevertheless, the accompanying chart, Table IV-1, list many of the factors that have to be assigned numerical values.)

Two Economic Risk Models Applied to Alternatives to Whole-Animal Testing

A search of the literature has not revealed many economic studies that have dealt specifically with alternatives to whole-animal testing for mutagenicity and carcinogenicity. Those that have been found have approached the problem from a general viewpoint and fall short of adequate quantification of the parameters.

The lack of quantification is understandable. Scientific assessment of the validity of short-term in vitro studies is still in question. The data base on the specific test or battery of tests that is recommended for any particular substance is weak and subject to the professional scientific judgment of the investigator. The economic production function including costs and prices is widely variable among different laboratories. The socital risks of misclassifying chemical substances are widely variable depending on a

TABLE IV-1

HIERARCHY OF RISK CONSEQUENCES AND SOME ILLUSTRATIVE SCALES OF MANAGEMENT

	Individuals		Collectives	
	Primary	Derived	Primary	Derived
1. Premature Death (Actual Death)	Perceived value of avoidance of: Immediate death Time Discounted death Statistically discounted death	Measured value of all 3. How much one would give to avoid these in terms of dollars, activities, & change of life style	Average no. of years foregone. No. of deaths per million per year. Cause of death.	Cost to society. Cost to average individuals. Dollars spent per death avoided (or years saved).
2. Illness & Disability (Pain & suffering, loss of faculties & social capability which may or may not lead to death)	Avoidance of severe, lasting pain & suffering. Dehabilitation. Death coupled with the above measures.	Amt. pd. for correction. Amt. pd. for prevention. Time & motion lost. Earnings foregone. Social activities foregone.	No. of illnesses & injuries. Purity of genetic pool.	Lost productivity. Cost of preventive & corrective health care.
3. Survival Factors	Adequacy of diet, shelter, & comfort. Crowding Factors.	Calories/day. Amt. per period. Portion of income for sustenance. People/one mile.	Avg. diet, shelter, etc. levels. No. & % with inadequate diet, shelter, etc. Birth rate/death rate.	Cost of adequate survival factors. Cost of avoiding deficiencies. Resultant ill health & death.
4. Exhaustible Resources	Unavailability to meet current & future needs. Cost to meet needs vs substitution.	Price - demand relationships. Substitutability.	Scarcity of resources. Dependence on imports. Price-demand relationships. Substitution elasticity.	Cost of waste & conservation. National offense vs industry needs. Import-export balance.
5. Security	Perception of rights encroached. Crime incidence as an individual. Sterility of Govt.	Measures of these.	Civil rights actions. No. of rights threatened, e.g., privacy, fair hearing, etc. Crime rate. Inflation rate.	Public opinion polls. Value of the dollar. Public safety dollars spent.
6. Belonging - Love	Perception of group. Family & society membership. Loneliness.	Measures of these.	Divorce rate. Marriage rate. Changing social mores.	Number & kinds of clubs & activities. Membership statistics.
7. Egocentric	Perception of wealth, power, rank, fame, etc. Material wealth. Freedom as an individual. Self esteem.	Earnings & assets. Position.	No. & kinds of people who have satisfied their personal goals.	Suicide rate.
8. Self-Actualization	Perception of the many factors of quality of life. Self-satisfaction.	Measures of the quality of life.	Collective measures of quality of life. Collective satisfaction.	Collective indices of the quality of life.

Source; W,D, Rowe, An Anatomy of Risk (43) number of factors. In short, there are enough uncertainties involved that the state-of-the-art in the evaluation of benefits and risks is so subjective that only the most general of models can be applied to the problem.

Nevertheless, the analyses that have been completed are instructive. The Lave, Omenn, Heffernan and Dranoff study (23) uses a cost-effectiveness approach. They look for a point where the cost of performing an extra test in a battery of in vitro tests offsets the benefits from increased accuracy. Since the degree of accuracy in unknown (even though they measure it by comparing the in vitro test results with whole-animal testing results), they attempt to find the battery of tests which will minimize social costs, defined by them as the sum of testing results and misclassification costs.

Testing costs per chemical substance for in vitro methods are a small percentage of the costs of R&D and of marketing the substance. Misclassification costs can be very large, depending on the particular chemical's potency and exposure levels in society. Since there are over 50,000 untested chemicals now in use and many more being put on the market each year, the authors apply "average" costs to misclassification errors. Although they are aware of the dangers in overgeneralizing with averages, they do not make an effort to adequately explain their choice of $1 million for a false positive and $10 million for a false negative error.

It is likely that their choice of economic costs for misclassification is on the conservative, low side. It could also be argued that the discrepancy between the values for false negatives and false positives of one order of magnitude is too great. On the surface, the false positive finding would affect the companies more than individuals, as it would prevent a safe chemical from being used by the public. Corporate profits would be affected as might the employment of some workers. The hidden cost is that the drug or chemical might be very beneficial to a group of people. By not permitting its manufacture and sale, hidden and unknown benefits could be lost. Much, of course, depends on the types of substitutes in the market for that product--their costs, risks, and availability. The important point is that false positives could be just as costly to society in given situations as false negatives, and the study should not be interpreted to mean that only false negatives are important to avoid.

By doing a sensitivity analysis on their data, the authors show that risks to society of not doing the short-term tests can be very large. If one agrees with their assumption concerning the relative social costs of false negatives with false positives, then this has implications for the cut-off point for determining which tested chemicals are carcinogens. Since it is currently impossible to test all chemicals, even using short-term tests, they argue that those with the highest use and potency should be tested first--certainly a logical conclusion for which an involved economic study is unnecessary.

Basically, they argue that it is cheaper for society to test all chemicals with a battery of short-term tests even though there will be misclassification errors. As a screening device, the potential social cost savings are large, given that the testing is inexpensive. As human knowledge of the reliability of the test increases from doing large numbers of these tests, they eventually may substitute for whole-animal tests.

There is a great deal of discussion concerning the cut-off points and decision points that should be considered in testing and regulation of chemicals. But their criteria are based on a very inadequate data set--38 chemicals out of over 50,000. The criteria do not include a discussion of the relative cost of whole-animal tests. Even though short-term test costs are extremely low relative to whole-animal chronic testing, given the very large potential social costs to misclassifying substances, whole-animal testing might still pass economic efficiency tests. And this is particularly true if initial screening has determined the most potent chemicals and the ones most likely to be used in large quantities.

The second study appears in Appendix E of the National Academy of Sciences Report, Toxicity Testing (34). This is a very interesting analysis of the regulatory and social costs of misclassification. It is conducted as a theoretical exercise--no real measurements are attempted. The output is a series of matrices taking into consideration different exposure levels and different toxicity levels.

The conclusions find some regularities to the structure of regulatory costs. They find that the costs to underregulate a chemical is much greater than the costs associated with overregulation. This corresponds to Lave's assumption that the social costs for false negatives is greater than that for false positives, even though in either case, specific examples could counter the generalization.

Second, they find that classifying the degree of exposure is important but not symmetric with the degree of toxicity of a substance. The toxicity factor plays a greater role in the social cost and risk analysis. The reason attributed to this is that exposure tends to be proportional to benefits, while the toxic potency is an inherent property of the chemical and does not vary with production, benefit, or exposure. They point out that toxic potency is determined by scientific investigation and is unlikely to change very much, but the data on production and market changes is dynamic and is subject to greater data inaccuracies and variabilities.

Both analyses are only as accurate as their underlying assumptions and data. Lave uses a variety of current data while the NAS study is based on theoretical probabilities. The Lave study specifically is aimed to show the economic efficiency of short-term testing procedures. The NAS study is more general--it attempts to assess the economic properties of any misclassification of chemical substances, whatever the source.

These analyses are important contributions to the literature. The frameworks that they set up for evaluation will be even more useful as additional data becomes available on both the scientific and economic aspects of alternatives to whole-animal testing. Their general conclusions about regulatory costs should give policy makers some guidelines for screening chemicals with the greatest potential for harm. But, the models are still too unspecific for reaching actual decisions in a political environment.

CONCLUSIONS AND RECOMMENDATIONS

Economic analyses can be very useful to policy and decision makers, but they must be used properly. All too often economic models and economists are criticized for inaccurate results or for poor predictions. Frequently, the cause of the problem lies not with the technical economic work but with the improper use of the results. It is not easy to understand all of the technical assumptions and jargon that accompany economic studies. Economic studies are only one of the many inputs policy and decision makers must consider when faced with problems to solve. The decision maker rarely has time to read studies in their entirety and to weigh the results of the analyses by all of the caveats and assumptions. Results tend to get simplified and used in ways that were never intended.

The thrust of this section has been to make the reader aware of the many limitations of economic analysis when applied to the very special case of regulating new developments in biochemistry that are on the threshold of possibly changing the way research and testing for mutagenicity and carcinogenicity is being done.

Economic thinking applied to this new technology is able to be of great help in setting up frameworks for analysis, in helping decision makers be aware of the types of consequences that might occur in society, and in providing a basis for future studies.

But the quantitative economic results that may be presented along with the frameworks must be read in a very different light. They are only indicative of the future, not determinative. Every author has emphasized when conducting these types of studies that the results are dependent on a number of assumptions that may not be representative of the universe or may contain large error terms. They are, at the current time, the author's best guess. Someone else's best guess could be quite different. At most, these quantitative studies should be looked at through a sensitivity analysis--what is the range of probable values that reasonable people would agree on.

At every stage of the development of new technologies, economic analyses are appropriate and should be funded. Early in the development, when most results are scientific in nature, the most one can conclude from economic studies is the probability of risk and benefit that could occur with different market outcomes. This is useful in business planning as well as in social decisions. Business can set strategies to fine-tune their development plans to capture new markets and to increae profits. Government can begin to look at social risks and benefits from new technologies and plan for action at appropriate times (not too soon to create disincentives in the market place and not too late to prevent possible injury or harm to individuals).

As the technology moves into an applied stage, where specific market potential is apparent and the testing and planning takes on product characteristics, there may be the beginnings of test data available. At this point, quantitative economic measures can be applied, with much caution and speculation. The industry is likely to be changing rapidly. The danger in funding economic analysis lies both in not doing enough studies and in not doing them frequently enough. Government in particular is likely to fund one or two studies that may take a year or so to complete. By the time the study is published, it is outdated. Unfortunately, the tendancy is to use those results and not fund additional analyses.

This can be very dangerous and lead to serious regulatory mistakes. Economic conditions are dynamic and change as rapidly as do new technologies in their formative stages. If one begins to measure social costs and benefits and to use those results in policy making, the potential for large errors is far greater when economic analyses are not redone at least as frequently as the technological change parameters.

For example, the Lave analysis described above was conducted during 1983. The battery of tests that they used was, at the time, considered state-of-the-art. As the data bank on testing is continuously being enlarged and updated, we now understand that some of those tests have been dropped in favor of others. Yet, it is likely that their economic results will be quoted for some time to come, as their study is one of a few in the field.

When a technology is used on a large scale in the stream of commerce, economic analyses take on a very different character. Much more detail is possible on the quantitative side. Historical reviews of that technology's impact can be ascertained, and the policy maker is not limited to making analogies to the development of similar technologies. Predictions of risks, costs, and benefits are based on more substantial data. The impact of regulatory changes can be more accurately assessed. The problem is that at this point, the major benefits or costs are history. Regulation then follows and may impact future generations, but it is then too late to correct past mistakes.

And, it should be emphasized that the results of even the best quantitative studies can be attacked since data are taken from samples, methodologies are often not fully adequate to explain results, and assumptions concerning the state of the economy underly the models.

Economic modelling and studies are useful. They should be funded and conducted routinely for any public effort that will have a significant impact. They should be used correctly, with the decision maker having full knowledge of their usefulness as well as their problems. Particularly in a science or technology policy context, economic models have a place that often is not well recognized by scientists or engineers. (Business and corporate planners have always conducted economic and market analyses of new technologies, but, as we discussed above, their motivation and desired results are far different and more easily defined and measured than those necessary for quantified social costs and benefits.)

Finally, the most important consideration for a public official in using the results of economic studies is that they are only one of many inputs into the decision making process. The results of these economic studies have to be evaluated not only on their technical and quantitative merits, but also along with the many other factors that go into decision making. Consequently, it should be virtually impossible to predict, from an economic context alone, the outcome of any individual policy decision regarding new technologies. But over a long period of time, economic factors will determine the eventual market acceptance and governmental acceptance of new technologies.

Whole-animal testing for mutagenicity and carcinogenicity is the current scientific standard. As short-term tests are conducted and more is learned about the accuracy associated with these short-term tests, economic motivations will force a much more careful look at the potential for regulatory acceptance of the in vitro test procedures. Resource cost efficiencies coupled with scientific developments most likely will push for their acceptance. Predicting when this will happen is impossible. Economics can only predict how large the social benefits and costs will be at any given stage in the process. Delay in changing regulations to accommodate new technologies can have large costs, just as premature decisions can increase societal risk exposure. It is up to the political process to make the decision.

REFERENCES/SOURCES

1. 1985 Budget of the U.S. Government., U.S. Government Printing Office.

2. Alternative Methods Could Cut Animal Use in Toxicity Tests, News Focus, *C&EN*, Wash., D.C., Oct. 31, 1983.

3. Arrow, Kenneth, J., "Economic Welfare and the Allocation of Resources for Invention" in "The Rate and Direction of Inventive Activity: Econ. Factors," NBER, Princeton University Press, Princeton, N. J. 1962.

4. Auletta, A. Genetic Toxicologist, U.S. EPA (personal communication, 1984.

5. Balthasar, H. U., Boschi, R., Menke, M., "Calling the Long Shots in R&D," SRI, International, Menlo Park, CA. 1978.

6. Bartsch, L., C. Malaveille, A. M. Camus, G. Martel-Planche, G. Brun, A. Huntefeuille, N. Sabadie, A. Barbin,T. Kuroki, C. Drevon, C. Piccoli and R. Montesano. 1980. Validation and Comparative Studies on 180 Chemicals with *S. typhimurium* strains and V79 Chinese hamster cells in the presence of various metabolizing systems. Mutation Res. 76:1-50.

7. Benson, S., Manager of Customer Services, Litton Bionetics (personal communication), 1984.

8. Brusick, D. 1983. Mutagenicity and Carcinogenicity Correlations Between Bacteria and Rodents. Ann. N.Y. Acad. Sci., Cellular Systems for Toxicity Testing: 164:176.

9. Brusick, D. 1983. The Use of Short-Term In Vitro and Submammalian Tests as Alternatives to Large Scale Animal Bioassays. Annals of The New York Academy of Sciences.

10. Brusick, D. 1983. Evaluation of Chronic Rodent Bioassays and Ames Assay Tests as Accurate Models for Predicting Human Carcinogens. In Application of Biological Markets to Carcinogen Testing. H. Milman and S. Sell, Ed.:153-163. Plenum Publ., New York, New York.

11. Brusick, D. Director, Molecular Sciences Directorate, Litton Bionetics (personal communication), 1984.

12. Cost Analysis Methodology and Protocol Estimates, TSCA Health Standards and FIFRA Guidelines, Study conducted by Enviro Control, Inc., Rockville, Maryland for the Office of Regulatory Analysis, U.S. EPA, Washington, D.C., April 1983.

13. Environmental Protection Agency, 1984. Chemical Carcinogens: A Review of the Science and Its Associated Principle. EPA of Science and Technology Policy. Draft Report: March 15, 1984.

14. Environmental Protection Agency, 1984. 40 CFR Part 158, Data Requirements for Pesticide Registration; Final Rule. Fed. Reg. 49(207):42856-42905.

15. Flamm, Gary, U.S. Food and Drug Administration (personal communication), 1984.

16. Genetic Schedule, Schedule of Prices, Bioassays Systems Corporation, 5/83

17. Genetic Toxicology Assays and Pricing Schedule, SRI International, March 1984.

18. Genetic Toxicology Fee Schedule, Hazelton Biotechnologies Corporation, 1/84
Acute Toxicology Fee Schedule, 1/84
Medical Devide and Biomaterial Testing Fee Schedule, 4/84
Genetic Toxicology, Schedule of

19. Gold, Bela, "Research, Technological Change, and Economic Analysis," in "The Quarterly Review and Economics and Business," Spring 1977.

20. Hazelton Industries, Annual Report for Fiscal Year 1984.

21. Kochanowski, P., Hertzfeld, H., "Often Overlooked Factors in Measuring the Return to Government R&D" in "Policy Analysis," Vol. 7, Number 2, Spring 1981.

22. Lave, L., Omenn, G., Heffernan, K., Dranoff, G. 1983. A Model for Selecting Short-Term Tests of Carcinogenicity. Journal of the American College of Toxicology. Vol. 2, No. 1.

23. Lave, L., Omenn, G., Heffernan, K., Dranoff, G. Cost Effectiveness of Short-Term Tests of Carcinogenicity. Manuscript, 1983.

24. Library of Congress, Congressional Research Service, U.S. Congress, "Cost-Benefit Analysis in Federal Regulations: A Review and Analysis of Development," 1978-1984, Report No. 84-74E, May 15, 1984.

25. Litton Bionetics, Inc. Price List for Molecular Toxicology Assays, October 1984.

26. Maclaurin, W. Rupert, "The Sequence from Invention to Innovation and Economic Growth," Quarterly Journal of Economics, February 1953, p.97.

27. Management of Laboratory Animal Facilities" Trends in Pharmacological Sciences, May 1983.

28. Mansfield, Edwin, "Microeconomics Theory and Applications," 4th Edition, W. W. Norton and Co., New York, 1982.

29. Mansfield, Edwin, "Contribution of R&D to Economic Growth in the U.S.," Science, Volume 175, Number 4021, February 4, 1972, p. 477.

30. McCann, J., E. Choi, E. Yamasaki and B. N. Ames. 1975. Detection of Carcinogens as Mutagens in the Salmonella/microsome tests: Assay of 300 Chemicals. Proc. Nat. Acad. Sci. USA 72:51356-5139.

31. McCann, J., Horn, L., and J. Kaldor. 1984. An Evaluation of Salmonella (Ames) Test Data in the Published Literature: Application of Statistical Procedures and Analysis of Mutagenic Potency. Mutation Res. 134:1-47.

32. McKean, Roland N., "The Nature of Cost-Benefit Analysis," an article in "Microeconomics," edited by E. Mansfield, W. W. Norton and Co., New York, 1979, p. 509.

33. Mogee, Mary E., "Time Lags in the Production of Technological Innovations," National Science Foundation, Washington, D.C., January 1976.

34. National Academy of Science, National Research Council. 1984. "Toxicity Testing, Strategies to Determine Needs and Priorities." Nat. Acad. Press, Washington, D.C., 1984.

35. National Science Foundation, Science Indicators, 1982, Appendix, Tables 4-7, 4-8; R&D Expenditures by industry, 1960-81.

36. National Toxicology Program, Board of Scientific Counselors, 1984. Draft Report - Ad Hoc Panel on Chemical Carcinogenesis Testing and Evaluation. J. Doull, Chm. Feb. 15, 1984.

37. Nelson, Richard R., "In Search of a Useful Theory of Innovation," in "Research Policy," 1977, p. 37.

38. Nelson, Richard, R., "The Simple Economics of Basic Scientific Research" in the "Journal of Political Economy," June 1959.

39. Office of Technology Assessment, U.S. Congress. "Assessment of Technologies for Determining Cancer Risks from the Environment." 1981.

40. Panem, Sandra. U.S. EPA (personal communication), 1984.

41. Rinkus, S.J. and M.S. Legator. 1979. Chemical characterization of 465 known or suspected carcinogens and their correlation with mutagenic activity in the Salmonella typhimurium system. Cancer Res. 39:3289-3318.

42. Rowan, A. N. and A. M. Goldberg. 1985. Perspectives on alternatives to current animal testing techniques in preclinical toxicology. Ann. Rev. Pharmacol. (in press)

43. Rowe, W. D., "An Anatomy of Risk," Environmental Protection Agency, Washington, D.C., March 1975.

44. Salsburg, D. 1983. The lifetime feeding study in mice and rats - an examination of its validity as a bioassay for human carcinogens. Fund. Appl. Toxicol. 3:63-67.

45. Salsburg, D. 1983. Letter. Fund. Appl. Toxicol. 3:6a-7a.

46. Squire, R. November 1981. Ranking Animal Carcinogens: A Proposed Regulatory Approach. Vol. 214, p. 877.

47. Stock Market Survey for Hazelton Labs, by Silverberg, Rosenthal and Co., April 9, 1984. Quoted by Henry Spira in phone conversation.

48. Testing Capability and Resources Department of Molecular Toxicology, Litton Bionetics, March 1984 (includes price list).

49. The Shift to Toxicology Tests in vitro". Chemical Week, May 26, 1982. p. 26.

50. Toxicology Research Services, Pharmakon Research International, Inc. Fee Schedule, 8/1/84.

51. U.S. Census of Manufacturers, Preliminary Statistics. 1982.

52. U.S. General Accounting Office, U.S. Congress, "Assessing the Output of Federal Commercially Directed R&D," PAD 79-69, Washington, D.C., August 27, 1979.

53. U.S. Industrial Outlook. 1984. U.S. Department of Commerce, January 1984.

54. Water, Michael. U.S. EPA (personal communication), 1984.

55. van der Lee, C., Akzo, N. 1983. A Comparison of Test Costs and Reporting Times Reported by Twenty-Five Contract Laboratories for Toxicological Research. Arnhem, The Netherlands.